UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1995

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 7B

ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major David Hart

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Gwendolyn Smith

Program Administrator, RDL
Johnetta Thompson

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1995

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-98-

0818

Public reporting burden for this collection of information is estimated to average 1 hour per response, including ... athering and maintaining the data needed, and completing and reviewing the collection of information. Send comme ... ction of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate ... y, Suite 1204, Arlington, VA 22202-4302, and to the Office of management and Budget, Paperwork Reduction Project (07

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December, 1995	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
USAF Summer Research Program - 1995 Graduate Student Research Program Final Reports, Volume 7B, Armstrong Laboratory	
6. AUTHORS Gary Moore	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research and Development Labs, Culver City, CA	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NI 4040 Fairfax Dr, Suite 500 Arlington, VA 22203-1613	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Contract Number: F49620-93-C-0063

12a. DISTRIBUTION AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for Public Release	

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Graduate Student Research Program (USAF- GSRP) is designed to introduce university, college, and technical institute graduate students to Air Force research. This is accomplished by the graduate students being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory Technical Directorates and Air Force Air Logistics Centers. Each participant provided a report of their research, and these reports are consolidated into this annual report.

14. SUBJECT TERMS	15. NUMBER OF PAGES
AIR FORCE RESEARCH, AIR FORCE, ENGINEERING, LABORATORIES, REPORTS, SUMMER, UNIVERSITIES	
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 7 is bound in two parts, 7A and 7B. Volume 7A contains #1-18, and Volume 7B contains reports #19-34. The Table of Contents for Volume 6 is included in both parts.

This document is one of a set of 16 volumes describing the 1995 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A, 5B, & 5C	Wright Laboratory
6A & 6B	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

GSRP FINAL REPORT TABLE OF CONTENTS i-xiv

FACILITATING TEAM INNOVATION: THE EFFECTS OF PERCEIVED ORGANIZATIONAL SUPPORT, IDEA GENERATION TECHNIQUE, AND PERSONALITY FACTORS

Steven C. Morgan
Graduate Student
Department of Psychology

Virginia Polytechnic Institute and State University
Blacksburg, VA 24060-0436

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling, Air Force Base, DC

and

Armstrong Laboratory

August 1995

FACILITATING TEAM INNOVATION: THE EFFECTS OF PERCEIVED ORGANIZATIONAL SUPPORT, IDEA GENERATION TECHNIQUE, AND PERSONALITY FACTORS

Steven C. Morgan
Graduate Student
Department of Psychology
Virginia Polytechnic Institute and State University

Abstract

The proposed study is part of a larger five-year R & D initiative to investigate innovation and flexibility in the workforce*. The current study will examine the effects of perceived organizational support, idea generation technique, and personality factors (social anxiousness, adaption/innovation, and need for cognition) on team innovation. To enhance generalizability of results, intact work teams from a large organization will serve as participants. One goal of this study is to address the discrepancy between the innovation literature and brainstorming research with regard to the optimal idea generation technique as the initial stage of innovation. Of further interest are the effects of satisfaction with the innovation process on idea generation. Idea generation technique will also serve as a moderator between the various personality factors and innovation in order to achieve a more comprehensive view of the innovation process.

* For more information on the five-year research program contact Dr. Sherrie P. Gott, senior scientist at: AL/HRMD, Manpower and Personnel Research Division, 7909 Lindberg Drive, Brooks AFB, Texas, 78235-5352

FACILITATING TEAM INNOVATION: THE EFFECTS OF PERCEIVED ORGANIZATIONAL SUPPORT, IDEA GENERATION TECHNIQUE, AND PERSONALITY FACTORS

Steven C. Morgan

Introduction

In order to be competitive in a rapidly changing environment, organizations need to be able to adapt to the dynamic economic, social, and political factors present in the system in which they operate (West & Wallace, 1991). One method by which organizations address this need is by innovating. Because the pace and complexity of environmental change limits the effectiveness of individual innovation, organizations are turning to teams as potential agents for change and innovation. While many organizations (e.g., Xerox, Motorola, GE) and researchers (e.g., Garvin, 1993; Cannon-Bowers, Oser, & Flanagan, 1992) advocate the use of teams as agents for innovation, little empirical research exists which examines the process of generating innovative ideas within a team context. In this report, the concepts of innovation, organizational support for innovation, and idea generation, as well as selected personality measures, will be treated descriptively. Next, an empirical study is proposed to investigate the idea generation stage of innovation among teams confronted with realistic, technical tasks.

Innovation

Process. The innovation process has been defined as: the intentional *introduction* and *application*, within a group or organization, of ideas, processes, products, etc., new to the adopting agent and designed to benefit the organization (West & Wallace, 1991); the *development* and *implementation* of new ideas by individuals who, over time, engage in transactions within an institution (Van de Ven, 1986); and, the successful *implementation* of creative ideas *generated* within an organization (Amabile, 1988). As evidenced by the preceding definitions, there is a general consensus in the literature that the innovation process consists of two major phases: idea generation and idea implementation.

Others offer a more comprehensive view of the innovation process. For example, Van de Ven (1986) delineates five stages of the innovation process: idea appreciation; articulation; adoption; institutionalization; and, decay. Kanter (1988) posits that there are four major tasks in the process of innovation: idea generation; coalition building; idea realization; and, transfer or

diffusion. Kanter suggests that these tasks are carried out at the micro-level by individuals and teams within an organization, with these micro processes facilitated or impeded by macro-level conditions. Ancona and Caldwell (1990) propose that teams follow the following pattern of innovation: creation of ideas; development of ideas; and, diffusion/exportation of ideas. Rosen and Servo (1991) postulate the following formula for innovation: innovation = conception + invention + exploitation. Regardless of which definition, process model, or conceptualization of innovation one espouses, an omnipresent theme throughout is that the starting point of the innovation process is the initial innovative idea.

Regarding team innovation in organizations, it is the product of multiple factors yet is always initiated by the creativity of individual members. As such, innovation is triggered by an individual's recognition of a new opportunity (i.e., Van de Ven's "appreciation"). These individuals or "idea generators" (Galbraith, 1982), consequently, initiate the process of departing from the organization's established routines (i.e., innovation) (Kanter, 1988). There are, however, differing perspectives in the literature regarding "how" innovation happens.

As emphasized by the innovation literature in general and Van de Ven (1986) in particular, innovation is not an individual process - it is a collective achievement. Once an idea is presented (i.e., "articulation"), over time a wide array of people with diverse perspectives become involved in the innovation process. These diverse perspectives, in turn, facilitate the proliferation of the one idea into multiple ideas. Hence, although an idea may be voiced by a single individual, the idea is actually a product of previous transactions with other members. Thus, within a team context, ideas are born, suggested, discussed, and fine-tuned so that they evolve into *team ideas*. These ideas are then exported for potential organizational implementation.

In short, this traditional view of the innovation process assumes that the building block of innovation, the innovative idea, will somehow automatically surface, evolve, and flourish in a team context. However, as will be discussed shortly, this assumption is suspect. Kanter (1988) suggests that micro processes such as idea generation are often affected by macro conditions within the organization (e.g., organizational support) as well as micro conditions such as personality characteristics of team members (e.g., social anxiousness) (Driskell, Hogan, & Salas, 1987).

Organizational Support. One macro-level condition that may affect idea generation is organizational support of innovation. For, even if individuals are capable of innovation, their willingness to engage in the process may be tempered by perceptions regarding the consequences of such actions in a given environment (Mumford & Gustafson, 1988). Organizations that show disinterest towards innovation or that over-emphasize maintaining the status quo create a climate in which the evolution of creative ideas may be suppressed. If it is perceived that an organization does not encourage innovation among employees or is apathetic towards any innovative ideas, why should employees and their teams put forth the effort? In the same vein, if an organization supports the maintenance of routine instead of supporting the change associated with innovation, idea generators may feel it is inappropriate to innovate. As Schein (1993) proposes, the organization must create a climate of psychological safety - employees must feel it is "safe" to be innovative. Schein suggests some ways this may be done: give employees opportunities for practice; let employees know it is okay to make mistakes; reward efforts in the direction of innovation; and, create norms that support innovation and experimentation. Other authors also stress the importance of organizational support for innovation (e.g., Henderson, 1994; Brown, 1991; Zuboff, 1991; Tyre & Orlinowski, 1993, Fleishman & Zaccaro, 1992).

Empirical studies provide added support for the influence of organizational culture on innovation. In a meta-analysis of the relationships between organizational innovation and its potential determinants, Damanpour (1991) found a statistically significant association between managerial support for change and innovation. Likewise, in a study of health care teams, West and Wallace (1991) found that cultural norms supporting diversity accounted for 42 per cent of the variance in team innovation, where innovation was quantified by the number of innovations (i.e., changes in work practices) introduced over a two year period.

Thus, both theoretical and empirical research support the notion that organizational support facilitates team innovation. However, to date, no research has directly examined the effects of perceived organizational support on both idea generation and satisfaction with the innovation process. The current study will address these issues.

It is also possible that by participating in an idea generation session, employees' perceptions of organizational support for innovation will increase. This is important in that the

actual act of idea generation, by way of increasing perceived support, may reinforce and facilitate future idea generation. Thus with each subsequent session, innovation may improve. To address this possibility perceived support for innovation will be measured both before and after the idea generation session.

Idea Generation Technique

As noted earlier, the initial innovative idea is the engine for a potential innovative practice. Brainstorming is widely recognized as a useful technique for idea generation in organizations. Idea generation in a team context typically takes one of two forms: interactive group or nominal group.

Interactive Group. Interactive brainstorming is the process of generating ideas in a group with an emphasis on generating as many ideas as possible. As put forth by Osborn (1957), the guidelines for this interactive process are: to be uncritical of others; to state any idea that comes to mind, no matter how wild; to aim for a large quantity of ideas; and, to build on the ideas of others. Osborn argues that the large quantity of ideas generated using the brainstorming process provide the catalyst to stimulate quality ideas. Those advocating this procedure presumed, based on principles of association, that others' generated ideas would stimulate novel associations for those exposed to the ideas. Specifically, the more ideas expressed, the greater the likelihood of new associations. Likewise, the greater the breadth of ideas expressed, the greater the likelihood of novel associations. These principles of association are in line with the aforementioned perspective on innovation that ideas naturally surface and thrive in a synergistic team context.

Nominal Group. The other form of brainstorming is known as the nominal group technique. Traditionally, it has been presupposed that interactive brainstorming is more efficient than individual idea generation (e.g., "more heads are better than one" adage). However, research comparing interactive brainstorming groups with similar members of individuals brainstorming alone (i.e., nominal groups) has demonstrated that nominal groups outperform their interactive counterparts (e.g., Diehl & Strobe, 1987; Diehl & Strobe, 1991; Paulus, Dzinonet, Poletes, & Comacho, 1993). Results generally indicate that nominal groups produce about twice as many ideas as interactive groups (see Mullen, Johnson, & Salas, 1991, for a review).

Possible explanations for the performance difference between interactive and nominal groups include: social psychological factors such as evaluation apprehension (concern for what others will think of ideas); procedural constraints such as production blocking (difficulty expressing ideas while others are talking or forgetting ideas while waiting for others to present theirs) (Diehl & Strobe, 1991); and economic factors such as social loafing (due to not being accountable for individual performance) (Shepperd, 1993). Whatever the explanation, the generally poor performance of interacting brainstorming groups presents an interesting challenge to the study of team innovation.

It should be noted, however, that some researchers (e.g., Paulus, Brown, & Ortega., in press) argue that existing evidence of the superiority of nominal techniques over interactive techniques is questionable because of the experimental tasks used to compare the two. As Paulus et al. (in press) point out, most studies comparing the two idea generation techniques used tasks of little personal relevance to the participants' lives (e.g., finding uses for an extra thumb). This provides a stark contrast to idea generation in organizations which involves a brainstorming task that has personal relevance (Paulus, Larey, & Ortega, 1995). Paulus et al. (in press) thus suggest that if a meaningful, relevant brainstorming task is employed, results may challenge those previously documented. To address this possibility the current study will be conducted in an organizational setting and will use a meaningful, relevant brainstorming task.

Because idea generation is the foundation of innovation, it is critical to determine which idea generation technique (interactive or nominal) is more effective in an organizational context. In other words, is it reasonable to assume that idea generation and innovation will naturally and optimally evolve in a team context (which amounts to an interactive technique)? Or, would the innovation process improve via the use of the nominal group technique (i.e., having team members generate ideas in isolation and then aggregating them)? The former technique assumes ideas stimulate other ideas to facilitate idea generation. The latter technique assumes the reduction of group inhibitory factors such as social loafing, blocking, and evaluation apprehension will facilitate idea generation as individuals work alone. In sum, facilitating team innovation begins with facilitating idea generation. The current study will examine which idea generation technique is best suited for influencing idea generation (innovation).

Of further interest to the current investigation is the impact of idea generation technique on satisfaction with the innovation process. Many researchers and practitioners believe that the reason interactive brainstorming continues to be widely used in organizations (despite the fact that it appears to be less efficient than nominal brainstorming) is that positive emotional responses (e.g., satisfaction) to brainstorming create perceptions of high team effectiveness despite low objective effectiveness. In other words, interactive teams rate their performance higher than nominal teams; however, nominal teams outperform the interactive teams. Indeed, the brainstorming literature indicates that, while nominal groups are better at idea generation, interactive teams consistently report that they are more satisfied with the process. R. I. Sutton (personal communication, August 3rd, 1995) offers a possible explanation for this phenomenon. He suggests that the "norms for good cheer" associated with traditional interactive brainstorming instructions (e.g., don't be critical of others; have fun developing wild ideas, etc.) may impede idea generation while at the same time foster satisfaction with the process.

However, the above relationship between satisfaction and idea generation documented in the brainstorming literature contradicts Isen's research (e.g., Estrada, Isen, & Young, 1994) on affect and creativity. Isen's work implies that if interactive contexts foster more positive emotion (greater satisfaction), teams will be *more* creative and innovative. Thus, another goal of the current study is to examine the effects of satisfaction on innovation to gain a better understanding of the innovative process and perhaps contribute evidence that bears on the debate surrounding the comparison of interactive and nominal techniques.

Personality Factors

Individual personality characteristics of group members may also play a key role in the idea generation phase of innovation (Driskell et al., 1987).

Adaption-Innovation. Kirton's (1976) Adaption-Innovation theory suggests that individuals have a cognitive style (i.e., preferred mode of tackling problems at various stages of the problem-solving process) which is independent of intellectual capacity and can be plotted on an adaption-innovation continuum. The adaptor is characterized as reliable, efficient, methodological, disciplined, and conforming. The innovator is characterized as undisciplined, tangential, and as approaching tasks from unsuspected angles (Kirton, 1987). Moreover, while

adaptors strive to maintain harmony in a team, innovators often promote conflict. Hence, adaptors are apt to be more influential in the idea implementation phase as they are more likely to be congenial. Innovators, on the other hand, may be more influential in the idea generation phase as they are more likely to challenge the status quo and voice innovative ideas.

Social Anxiousness. Social anxiousness represents the degree to which one feels anxiety and distress across social interactions (Leary, 1983). Social anxiousness has been found to be related to concerns about how one is perceived and evaluated by others (Leary & Kowalski, 1993). Recent research by DePaulo, Epstein, and LeMay (1990) found that high socially anxious individuals (classified via the Interaction Anxiousness Scale; Leary, 1983) who thought they were being evaluated were more inhibited in conversation. In terms of brainstorming, there is evidence that interactive groups low in social anxiousness are significantly more productive than high socially anxious groups (Camacho & Paulus, 1995).

Need for Cognition. The construct of need for cognition (NFC) is based on the proposition that there are individual differences in the predisposition towards mental laziness or mental engagement (Petty & Cacioppo, 1986). Thus, the need for cognition is the tendency for an individual to engage in and enjoy thinking. Evidence suggests that groups with lower NFC scores (as measured by the NFC Scale; Cacioppo & Petty, 1982) are less effective at group brainstorming than groups with higher average NFC scores.

While the aforementioned research on personality factors indicates they may affect innovation, no research to date has directly examined how the effects of personality on innovation and satisfaction with innovation may be moderated by the idea generation technique employed. The current study will address this by testing a path model in which idea generation technique will serve as a moderator between personality factors and innovation and satisfaction with innovation.

Hypotheses

Innovation Hypothesis 1. Consistent with priciples of association and group synergy and the innovation literature, teams and individuals will be more innovative in an interactive context than in a nominal team context.

Innovation Hypothesis 1b. As documented in the brainstorming literature, there will be a process loss associated with the group context. Thus, teams and individuals will be more innovative in a nominal team context than in an interactive team context.

Innovation Hypothesis 2. Teams and individuals which perceive more organizational support for innovation will be more innovative than those which perceive lower organizational support for innovation.

Innovation Hypothesis 3. There will be an ordinal interaction such that teams and individuals classified as innovative (via Kirton's KAI scale) will be more innovative than those identified as adaptive; however, because innovators are more conflict oriented, innovative teams and individuals will be more innovative in a nominal context than in an interactive context.

Innovation Hypothesis 4. Because social interaction may inhibit the participation of socially anxious individuals, idea generation technique will moderate the effects of social anxiousness on innovation. Thus, there will be a disordinal interaction with teams and individuals high in social anxiousness (classified via the IAS) being more innovative in the nominal context than in the interactive context while those low in social anxiousness will be more innovative in the interactive context than in the nominal context.

Innovation Hypothesis 5. Because those that are high in need for cognition have a predisposition towards mental engagement, teams and individuals identified as having a high need for cognition (classified via the NFC Scale) will be more innovative than those that have a low need for cognition.

Innovation Hypothesis 6. Consistent with the brainstorming literature, teams and individuals which report that they are less satisfied with the innovation process will be more innovative.

Innovation Hypothesis 6b. Research on affect and creativity suggests that a positive mood state leads to greater creativity. Thus, teams and individuals which are more satisfied with the innovation process will be more innovative.

Satisfaction with Innovation Hypothesis 1. Because team members often enjoy working in teams and have more favorable mood states than individual performers (Paulus & Dzindolet,

1993), teams and individuals in the interactive context will be more satisfied with the innovation process than those in the nominal context.

Satisfaction with Innovation Hypothesis 2. Because organizational cultural mores may influence perceptions about group productivity and teamwork (Paulus et al., in press), teams and individuals which perceive more organizational support for innovation will be more satisfied than those which perceive lower organizational support.

Satisfaction with Innovation Hypothesis 3. Because the cognitive style of innovators (classified via the KIA Scale) is more amenable to the idea generation phase, innovative teams and individuals will be more satisfied with the process than adaptors.

Satisfaction with Innovation Hypothesis 4. Because social interaction may cause anxiety and distress for socially anxious individuals, idea generation technique will moderate the effects of social anxiousness on satisfaction with innovation. There will be a disordinal interaction such that teams and individuals high in social anxiousness will be more satisfied with the innovation process in the nominal context as opposed to the interactive context. However, those low in social anxiousness will be more satisfied in the interactive context because of the positive experience of working as a team.

Satisfaction with Innovation Hypothesis 5. There will be an ordinal interaction such that teams and individuals classified as high in need for cognition (via the NFC Scale) will be more satisfied with the innovation process than those identified as low in need for cognition; however, because there is less pressure to cognate in a group setting (i.e., social loafing) those low in NFC will be more satisfied in the interactive context than in the nominal context.

Perceived Organizational Support Hypothesis 1. Perceptions of organizational support for innovation will increase after participating in an idea generation session.

Method

Participants

18 intact work teams (consisting of at least six team members) from a large organization will participate in the current investigation. The study will utilize a two-level (idea generation

technique) factorial design. For each team, individual members will be randomly assigned to conditions so that half the team will use the interactive technique while the other half will use the nominal technique. The gender composition of the team will not be controlled.

Experimental Task

Some researchers (e.g., Paulus, Brown, & Ortega., in press) argue that existing evidence of the superiority of nominal techniques over interactive techniques is questionable because of the experimental tasks used to compare the two. As Paulus et al. (in press) point out, most studies comparing the two idea generation techniques used laboratory designs in which ad hoc groups of college participants, who were virtual strangers, performed a task of little personal relevance to their lives (e.g., finding uses for an extra thumb). This provides a stark contrast to idea generation in organizations which involves group members who have worked together in other contexts and a brainstorming task that has personal relevance to them and the organization (Paulus, Larey, & Ortega, 1995). In laboratory designs there may be little need for interactive idea generation because individuals usually have an adequate pool of ideas to address the problem, thereby negating the need to draw on other members' knowledge. Conversely, in an organizational setting individuals may have a need for novel ideas and perceive others as providing needed stimulation for idea generation. Paulus et al. (in press) thus suggest that if a meaningful, relevant brainstorming task is employed, results may challenge those previously documented.

To address this possibility the current study will be conducted in an organizational setting and will use a meaningful, relevant brainstorming task. Through observation and the interviewing of team members, a brainstorming task will be devised deemed to be meaningful and relevant to a given team. All participants will perform the same brainstorming task in either the interactive group or nominal group context.

Procedure

Participants will be asked to complete a questionnaire before the experimental session. Each questionnaire will contain relevant demographic items, personality items (i.e., KAI; IAS; NFC), and items to assess perceived support for innovation (SSSI).

Each team will be comprised of three individuals. Each participant will be given a tape recorder and a personal microphone. Participants will then be given a copy of the brainstorming instructions and will be asked to read them along with the experimenter. The instructions will contain a statement of the problem and will remain with the participants during idea generation. The participants will be instructed to follow the rules of brainstorming as outlined by Osborn (1957): (a) don't be critical of others; (b) state any idea that comes to mind, no matter how wild; (c) aim for a large quantity of ideas; and, (d) build on the ideas of others. The participants will then be told that they have 15 minutes to brainstorm and that they are not to write anything down on paper or touch the tape recorders. The experimenter will then answer any questions participants have about the process. For the interactive condition, the experimenter will then turn on the recorders, tell the participants to begin, and leave the room. For the nominal conditions, the experimenter will place participants in separate rooms (each equipped with a tape recorder and microphone) so that each can generate ideas in isolation. After 15 minutes, the experimenter will return to the room(s) and turn off the recorders. Participants will then be asked to complete a questionnaire which will include items to assess perceived organizational support for innovation as well as items to assess satisfaction with the innovation (i.e., idea generation) process. Participants will then be asked to complete a debrief form and will be excused.

Exogenous Variables

Perceived Organizational Support. Perceived organizational support for innovation will be measured using the Siegel Scale of Support for Innovation (SSSI; Siegel & Kaemmerer, 1978). This 24 item scale is a shortened version of the original 61 item scale. These 24 items (e.g., "creativity is encouraged here") all load highly on a "support of creativity" factor identified by Siegel and Kaemmerer. This factor represents the extent to which organizational members perceive the organization as supporting independent functioning in pursuit of innovative ideas. It also includes the perception that the organization is open and adaptive to change. Participants will rate each item based on a six point Likert-type scale anchored by: 1 = "strongly agree" and 6 = "strongly disagree". This scale will serve as both a pre- and post-measure of perceived support. For team-level analyses the average team score on the SSSI will serve as an indicator of

perceived support. For individual-level analyses, the individual's SSSI score will serve as an indicator.

Personality Variables. Cognitive style (i.e., adaptation vs. innovation) will be assessed using Kirton's (1987) Adaptation-Innovation Inventory (KAI). This scale measures response to organizational change based on a scale ranging from an ability to "do things better" to the ability to "do things differently".

Social Anxiousness will be measured using the Interaction Anxiousness Scale (IAS; Leary, 1983). For each item (e.g., "I wish I had more confidence in social situations") of this 15 item scale participants will be asked to indicate the "degree to which the statement is characteristic or true of you". Participants will rate each item based on a five point graphic rating scale anchored by: 1 = "not at all characteristic"; 2 = "slightly characteristic"; 3 = "moderately characteristic"; 4 = "very characteristic"; 5 = "extremely characteristic".

Need for cognition (NFC) will be assessed using Cacioppo, Petty, & Kao's (1984) Need for Cognition Scale. Participants will be asked to indicate the degree of agreement or disagreement with each of 18 items (e.g., "I prefer my life be filled with puzzles that I must solve"). Items will be rated based on a nine point Likert-type scale anchored by: -4 = "very strong disagreement"; -3 = "strong disagreement"; -2 = "moderate disagreement"; -1 = "slight disagreement"; 0 = "neither agreement nor disagreement"; +1 = "slight agreement"; +2 = "moderate agreement"; +3 = "strong agreement"; +4 = "very strong agreement".

For team-level analyses the average team score on a given scale will be used as a measure for the respective personality variable. For individual-level analyses the individual scale score will be used.

Moderator Variable

Idea Generation Technique. Participants in the interactive groups will brainstorm interactively with other team members. Participants will be seated around a circular table with the chairs positioned so that all team members will be facing each other.

Participants in the nominal groups will brainstorm in isolation.

Endogenous Variables

Innovation. For purposes of the current investigation, innovation will be operationally defined as the number of unique ideas generated. Consistent with Osborn's presumption that the greater the quantity of ideas, the greater the quantity of high quality ideas; and, with the empirical evidence that quantity measures are highly correlated with quality measures (e.g., Diehl & Strobe, 1987; Dennis & Valacich, 1994), the number of ideas generated will also serve as a surrogate measure for quality of ideas. For team-level analyses innovation will be quantified both as the number of unique ideas generated by a team; and, the average number of unique ideas generated per person for a team. For individual-level analyses, innovation will be quantified as the number of unique ideas generated by a given individual. In both cases, expert raters will quantify innovation by analyzing the transcriptions of the tape recordings.

Satisfaction with Innovation. Satisfaction with innovation will be assessed using a measure based on items used by Dennis and Valacich (1992). Participants will be asked to indicate their level of satisfaction with each of five items (e.g., "How do you feel about the process by which you generated ideas?"). Items will be rated on a seven point scale anchored by: 1 = "Very dissatisfied" and 7 = "very satisfied". For team-level analyses the average team score on the scale will be used as a measure for satisfaction with innovation. For individual-level analyses the individual scale score will be used.

Analyses

As posited by Wright (1921; as cited in Pedhauser, 1982), in cases where causal relations are uncertain, the method of path analysis is optimally suited for finding the logical consequences of any particular hypothesis with regard to them. The current study will employ path analysis to analyze the proposed path model (see Figure 1) based on the various hypotheses related to

---

Insert Figure 1 about here

---

innovation and satisfaction with innovation. After assessing the goodness of fit of the path model, the proposed main effects (i.e., innovation hypotheses1, 1b, 2, 5, 6, 6b; satisfaction with innovation hypotheses 1, 2, 3) will be tested by testing the relevant path coefficient for significance. Moderating effects will be tested using moderated regression analyses with idea generation technique serving as a moderator (innovation hypotheses 3, 4; satisfaction with innovation hypotheses 4, 5). Perceived organizational support hypothesis 1 will be tested using repeated measures regression analysis, with pre- and post-SSSI scores serving as the repeated measures.

References

Amabile, T. M. (1988). A model of creativity and innovation in organizations. In B. M. Staw & L. L. Cummings (Eds.), Research in organizational behavior (Vol. 10) (pp. 123-167). Greenwich, CT: JAI Press.

Brown, J. S. (1991). Research that reinvents the corporation. Harvard Business Review, 69, 170.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. Journal of Personality and Social Psychology, 42 (1), 116-131.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. Journal of Personality Assessment, 48, 306-307.

Camacho, L. M., & Paulus, P. B. (1995). The role of social anxiousness in group brainstorming. Journal of Personality and Social Psychology, 68, 1071-1080.

Cannon-Bowers, J. A., Oser, R., & Flanagan, D. L. (1992). Work teams in industry: A selected review and proposed framework. In R. Swezey & E. Salas (Eds.), Teams: Their training and performance (pp. 355-377). Norwood, NJ: Ablex.

Damanpour, F. (1991). Organizational innovation: A meta-analysis of effects of determinants and moderators. Academy of Management Journal, 34 (3), 555-590.

Dennis, A. R., & Valacich, J. S. (1993). Computer brainstorms: more heads better than one. Journal of Applied Psychology, 78 (4), 531-537.

Dennis, A. R., & Valacich, J. S. (1994). Group, sub-group, and nominal group idea generation: New rules for a new Media? Journal of Mangement, 20 (4), 723-736.

DePaulo, B. M., Epstein, J. A., & LeMay, C. S. (1990). Responses of the socially anxious to the prospect of interpersonal evaluation. Journal of Personality, 58, 623-640.

Diehl, M., & Strobe, W. (1987). Productivity loss in brainstorming groups: The solution of a riddle. Journal of personality and Social Psychology, 53, 497-509.

Diehl, M., & Strobe, W. (1991). Productivity loss in idea-generating groups: Tracking down the blocking effect. Journal of Personality and Social Psychology, 61, 392-403.

Driskell, J. E., Hogan, R., & Salas, E. (1987). Personality and group performance. In C. Hendrick (Ed.), Group processes and intergroup relations: Review of personality and social psychology, Vol. 9 (pp 91-112). Newbury Park, CA: Sage.

Estrada, C. A., Isen, A. M., & Young M. J. (1994). Positive affect improves creative problem solving and influences reported source of practice satisfaction in physicians. Motivation and Emotion, 18, 4-15.

Fleishman, E. A., & Zaccaro, S. J. (1992). Toward a taxonomy of team performance functions. In R. Swezey & E. Salas (Eds.), Teams: Their training and performance (pp. 31-56). Norwood, NJ: Ablex.

Galbraith, J. (1982). Designing the innovating organization. Organizational Dynamics, 10, 5-25.

Garvin, D. A. (1993). Building a learning organization. Harvard Business Review, 71, 78-91.

Henderson, R. (1994). Managing innovation in the information age. Harvard Business Review, Jan-Feb, 100-105.

Kanter, R. M. (1988). When a thousand flowers bloom: Structural, collective, and social conditions for innovation in organization. In B. M. Staw & L. L. Cummings (Eds.), Research in organizational behavior (Vol. 10) (pp. 169-211). Greenwich, CT: JAI Press.

Kirton, M. J. (1976). Adaptors and innovators: A description and measure. Journal of Applied Psychology, 61, 622-629.

Kirton, M. J. (1987). Kirton adaption-innovation inventor manual (2nd ed.). Hatfield, England: Occupational Research Center.

Leary, M. R. (1983). Social anxiousness: The construct and its measurement. Journal of Personality Assessment, 47, 66-75.

Leary, M. R., & Kowalski, R. M. (1993). The interaction anxiousness scale: Construct and criterion-related validity. Journal of Personality Assessment, 61, 136-146.

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. Psychological Bulletin, 115, 210-227.

Mumford, M. D., & Gustafson, S. B. (1988). Creativity syndrome: Integration, application, and innovation. Psychological Bulletin, 103, 27-43.

Osborn, A. F. (1957). Applied imagination. New York: Scribner.

Paulus, P. B., Brown, V., & Ortega, A. H. (in press). Group cretivity. In R. E. Purser & A. Montuori (Eds.), Social creativity in organizations. Cresskill, NJ: Hampton Press.

Paulus, P. B., Dzindolet, M. T., Poletes, G. W., & Camacho, L. M. (1993). The perception of performance in group brainstorming: The illusion of group productivity. Personality and Social Psychology Bulletin, 19, 78-89.

Paulus, P. B., Larey, T. S., & Ortega, A. H. (1995). Performance and perceptions of brainstormers in an organizational setting. Basic and Applied Social Psychology, 17, 249-265.

Pedhauser, E. J. (1982). Multiple regression in behavioral research (2nd ed.). Orlando, FLA: Harcourt, Brace, Jovanovich.

Petty, R. E., & Cacioppo, J. T. (1986). Communication and persuasion: Central and peripheral routes to attitude change. New York: Springer-Verlag.

Rosenfeld, R. & Servo, J. C. (1991). Facilitating innovation in large organizations. In J. Henry & D. Walker (Eds.), Managing innovation (pp. 28-39). London, UK: Sage.

Schein, E. H. (1993). How can organizations learn faster? The challenge of entering the green room. Sloan Management Review, Winter, 85-92.

Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. Psychological Bulletin, 113, 67-81.

Siegel, S. M., & Kaemmerer, W. F. (1978). Measuring the perceived support for innovation in organizations. Journal of Applied Psychology, 63 (5), 555-562.

Tyre, M. J., & Orlinowski, W. J. (1993). Expoiting opportunities for technological improvement in organizations. Sloan Management Review, 13-26.

Van de Ven, A. H. (1986). Central problems in the management of innovation. Management Science, 32(5), 590-607.

West, M. A., & Wallace, M. (1991). Innovation in health care teams. European Journal of Social Psychology, 21, 303-315.

Wright, S. (1934). The method of path coefficients. Annals of Mathematical Statistics, 5, 161-215.

Zuboff, S. (1991). Can research reinvent the corporation? Harvard Business Review, 69, 164-165.

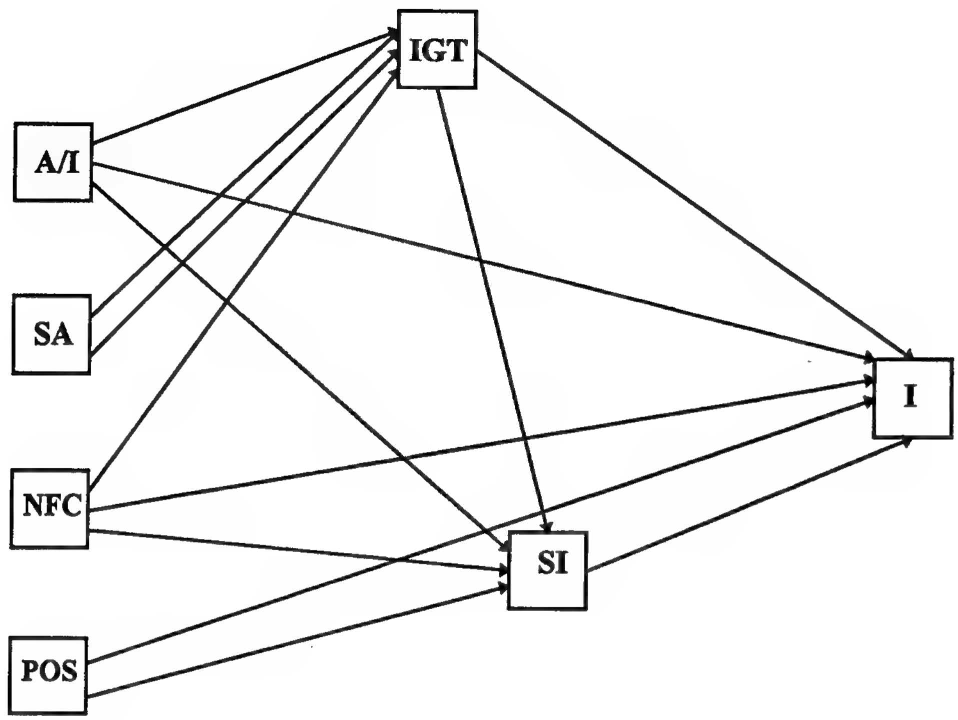


Figure 1. Path model for the effects of perceived organizational support for innovation (POS), idea generation technique (IGT), adaption/innovation (A/I), social anxiousness (SA), and need for cognition (NFC) on innovation (I) and satisfaction with the innovation process (SI). Note: paths leading to IGT represent a moderation effect not a direct effect (i.e., personality factors don't *cause* IGT).

THE ANALYSIS AND DISINFECTION OF DENTAL UNIT WATER LINES

Jennifer F. Nemeth
Graduate Student
Department of Chemistry

Emory University
1515 Pierce Dr.
Atlanta, GA 30322

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored By:
Air Force Office of Scientific Research
and
Armstrong Laboratory

August 1995

THE ANALYSIS AND DECONTAMINATION OF DENTAL UNIT WATER LINES

Jennifer F. Nemeth
Graduate Student
Department of Chemistry
Emory University

Abstract

The study was performed to determine the extent of microbial contamination in dental unit water lines at four U.S. Air Force Dental Clinics to evaluate a method for decontaminating and maintaining them on a routine basis. Experimental results indicated water from dental unit water lines in untreated units was contaminated with $>10^5$. Both 1:10 and 1:100 chlorine bleach (NaOCl) solutions were capable of reducing bacterial growth over time. A simple method for cleaning the units was developed using a 1:10 sodium hypochlorite solution. The preliminary evaluation suggests that a 1:100 dilution of sodium hypochlorite may also be effective for long term control of dental unit water line bio-film following initial disinfection.

Introduction

Over the past forty years, there has been increasing interest in the microbial contamination existing in dental unit water lines. Of the total bacterial contamination, only a small percentage originate from human oral flora and the remainder of the bacteria are environmental in origin; mainly gram negative water bacteria[1,2]. While proposed safe drinking water standards set a limit of 500 Colony Forming Units/mL (CFU/mL), most untreated dental units contained bacteria in excess of 10^5 CFU/mL[1,3,4,5].

Illnesses resulting from exposure to aerosols generated from the dental unit handpieces have been reported. A fatal case of Legionellosis in a dentist may have resulted from exposure to contaminated dental unit water[6]. Several investigators have isolated the Legionella, as well as Pseudomonas species, in contaminated dental water[7,8]. Concerns about the potential for infection, particularly in immunocompromised patients, has spurred a desire to prevent contamination in dental units. Methods currently under examination to accomplish this include flushing[9,10], filtration[11,12], chemical disinfection[4,13,14,15], and separate water reservoir systems[4,13,14].

The effectiveness of separate water systems, coupled with chemical disinfection with sodium hypochlorite solutions, was examined in this study. The first segment of the study was the analysis four U.S. Air Force dental clinics in San Antonio, Tx.

Four units at three clinics were equipped with ADEC separate reservoir systems (ADEC Inc., Newborg, OR) designed to isolate dental water supplies from municipal water. Each clinic had been treating units on a weekly basis with 1:10 sodium hypochlorite for periods ranging from six months to six years according to manufacturer's directions. Distilled water, prepared on the premises using commercial laboratory distillers, was used to fill the reservoirs for patient treatment. The remaining twelve units were located at a Lackland that had been connected to municipal water supplies for three years. The following chart summarizes the condition of the units at each base:

Air Force Base	Age of Unit	Time Untreated	Time Treated
Brooks AFB	8 yrs.	2 yrs.	6 yrs.
Kelly AFB	6 mon.	0 yrs.	6 mon.
Lackland AFB	3 yrs.	3 yrs.	0 yrs.
Randolph AFB	6 yr.	5 yrs.	1 yr.

The second segment to the study focused on the treatment of a set of dental units at Lackland AFB. The main focus in this analysis is to determine the composition of the water and water lines, the concentration of the sodium hypochlorite disinfectant required to control bio-film formation in dental unit water lines, and a method for periodic re-treatment of the units after they have been initially decontaminated.

Experimental

The procedure for part one of the study called for the analysis of four units at each of three bases: Brooks, Kelly, and Randolph as well as twelve units from Lackland for a total of twenty-four units. The additional units at Lackland were also used for the decontamination study. Water and line samples were taken from each of the units and analyzed for bacterial contamination and metal concentrations in out-put water and bio-film and inorganic deposits on tubing walls.

Water samples were obtained from all of the water lines; the number of lines depending on the model and the condition of the unit. A 200 mL pooled water sample was obtained with each line contributing approximately equal amounts. The samples were collected in sterile 250 mL sample bottles. The lines used were the assistant's and doctor's air-water-syringes and all the operable handpiece lines. Three main analyses were performed. 1) Inductive Coupled Plasma (ICP) was used to determine metal concentrations of Ni, Cr, Cu, Zn, and Ca in the water. The samples were prepared by taking 75-100 mL of water and preserving them with 0.1 mL of ultrapure, concentrated nitric acid. The Occupational, Environmental, and Analytical laboratory (AL/OEA) at Brooks AFB performed the ICP analysis. 2) Bacterial counts were determined using Millipore SPC Samplers (Millipore Inc., Marlborough, MA) which contain standard plate count media and provides a colony count after five days. Dilutions ranging from 1:10 to 1:10000 were used based on the of expected colonies following manufacturer's

instructions. All samples were incubated at 25°C to most accurately simulate the temperature of the water in the units and to promote the growth of slow-growing water bacteria which are characteristically recovered from dental unit water lines. The colonies were counted and the results reported in CFU/mL. 3) Acridine orange stain was used to confirm the presence of bacterial cells in the effluent water. The analysis was performed by spotting a microscope slide with a drop of sample water. The specimen was then dried, fixed with methanol, and stained with acridine orange which causes bacteria to fluoresce when viewed under ultraviolet light.

In addition to water analyses, the water lines were removed from the dental units for examination and water line segments that were thought to exhibit a representative degree of bio-film formation were chosen.[16] A line was selected and the outer casing disinfected with 70% isopropyl alcohol for one minute. A 10-15mm section of tubing was clamped with hemostats, and removed. Samples were cut from the inlet line of the control block, from the provider's air-water syringe, and from a handpiece line. The water that was trapped in the line between the clamps was subjected to the acridine orange analysis and three analyses were run on the tubing samples. 1) A 3-4mm aseptically retrieved piece of tubing was dropped into a culture tube (TC) containing Trypticase Soy Broth (TSB) and incubated at 25°C for five to seven days. The positive specimens were identified using standard bacteriological

methods to rule out external contamination from skin flora and confirm the presence of gram negative rods consistent with typical water line flora. 2) Light Microscopy (LM) was used to determine whether there were visible deposits on the lines. The samples were prepared by taking longitudinal and cross-sections of the lines to expose the interior. They were mounted onto studs with Colloidal Graphite (carbon paint) and examined at 25X to 200x magnification. 3) Upon completion of the LM analysis, the samples were coated with a Au/Pt layer for scanning electron microscopy (SEM) examination. The lines were examined by SEM for presence of bio-film, bacteria cells, or inorganic deposits. Analyses were run between 500x to 5000x.

The decontamination experiment used twelve dental units at Lackland AFB, which were connected to the San Antonio Water System for three years. After the initial microbial and chemical analysis described in part one, an additional resampling was performed on the day of the initial decontamination. After completing the initial sampling, the units were placed into one of three groups: 1:10 NaOCl (5000 ppm chlorine), 1:100 NaOCl (500 ppm chlorine), or a control. All units underwent a weekly treatment with an assigned solution. The controls were flushed with sterile deionized water (7.0 pH). Biological counts were monitored for the first week by sampling water from the units at 24, 48, 72 and 120 hours and immediately prior to the next disinfection procedure. The data was correlated with a chlorine analyses reported in a separate

report.[17] In addition to the water samples that were taken every week for SPC and ICP analysis, water line samples were taken at two and four weeks for LM and SEM observations.

In order to perform weekly flushing procedures, the dental units were removed from the city water system and ADEC Separate Reservoir Systems were installed. Once treatment began, only bottled, deionized sterile water was run through the units during patient treatment. The sodium hypochlorite that was used in the study was Clorox bleach (The Clorox Co., Oakland, CA). The experimental units were broken down into two groups of six units each with two controls, two 1:10 NaOCl, and two 1:100 NaOCl.

Prior to disinfecting the dental units, the water lines were purged with air to remove as much water as possible. The reservoir bottle was then filled with sterile bottled water, 1:10 bleach, or 1:100 bleach, depending on the unit, and the lines flushed for approximately thirty seconds to fill them with the cleaning solution. The lines were allowed to sit for approximately ten minutes with the solution in the them. The system was then air-purged, flushed with approximately 500 mL of bottled sterile water, air-purged again, and left dry until the next use. Although sterile water was left in the units following patient care, they were air purged each Friday and left dry until Monday.

Results

The bacteria counts at Brooks, Kelly, and Randolph ranged from <10 CFU/mL to >10^4. The Lackland units were all >10^5 CFU/mL as

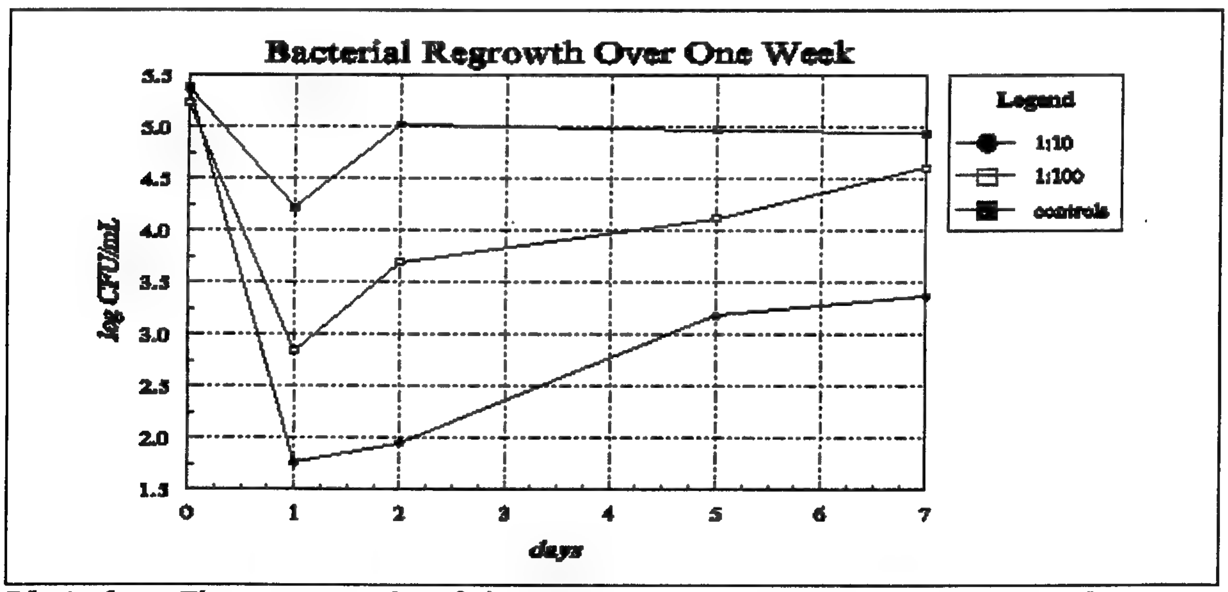
expected. The LM revealed deposits in many of the tubing samples from all of the units and the SEM identified which lines most probably had bio-film in them. The tubing cultures helped to confirm the presence of the bio-film. Most of the cultures at Brooks, Kelly, and Randolph exhibited negative results, while those at Lackland were mainly positive. As the experiment progressed, more culture became negative. The data for Brooks, Kelly, and Randolph can be found in Appendix B. Lackland's observational data can be found under the dates 6/20 and 7/5 in appendix D. No ICP data was reported for Brooks, Kelly, or Randolph.

The decontamination of the Lackland units decreased the bacterial counts from the initial 10^5 CFU/mL to counts of approximately 10^3 in the 1:10 and 1:100 NaOCl units and 10^4 in the controls (See Appendix C for complete data). There was an inverse relationship between the bacterial counts and the chlorine recovery from the units. The initial recovery was 86.9% for the 1:10 and 65.3% for the 1:100 NaOCl. After three weeks the percent recovery rose to 95.4% and 91.1% respectively.[17] A complete set of the data for Lackland can be found in Appendix D. The ICP data has yet to be completely correlated; however, there appears to be an increase in Ni with time. Initially, Ca was present in the water, but after the units were put on sterile deionized water the Ca fell below the detection limits. In the last week Ca was again detected in some of the samples (see appendix A for complete data).

Discussion

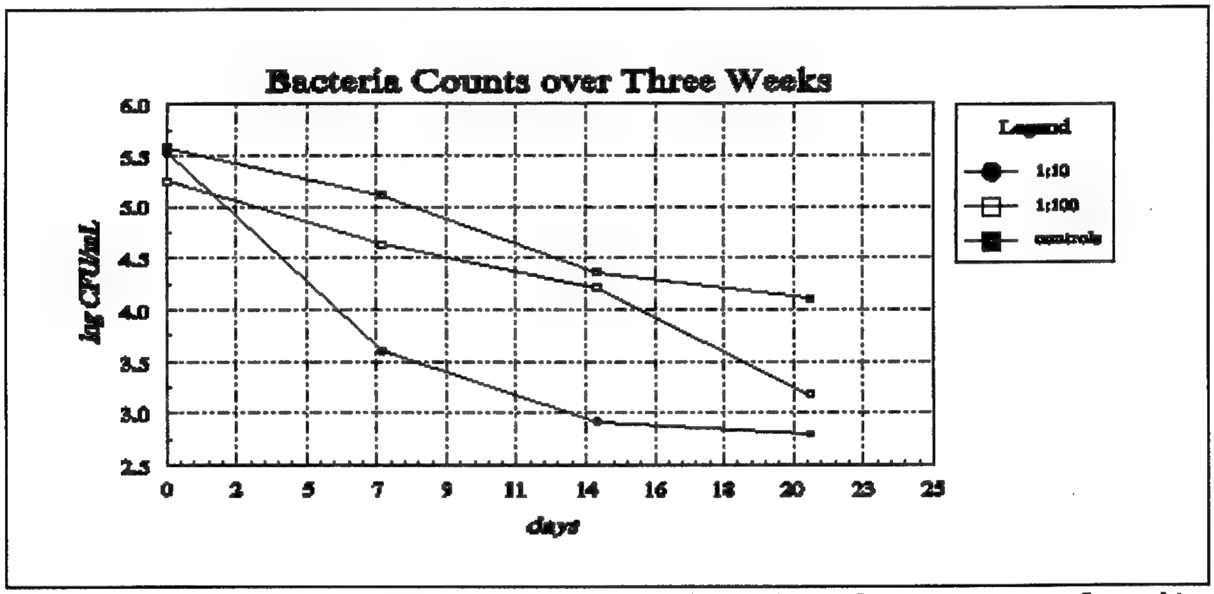
There was little consistency in the treatment of dental units by clinic personnel. All three bases on separate reservoir systems had at least one contaminated unit. Multiple factors probably contributed to this phenomenon, especially the technician's failure to treat lines that were not in use and not realizing that all of the water lines had to be cleaned whether or not coolant water was routinely used in them. Regrowth of bacteria was found to occur during the week between treatments. This quick regrowth suggests that it might be necessary to disinfect dental units twice a week instead of weekly.

The disinfection study revealed a number of results. First, cleaning with a bleach solution did not immediately eliminate all of the bio-film in the lines. Plot 1 illustrates the growth during the first week after decontamination. There was a decrease in the



Plot 1: The regrowth of bacteria after the first week of disinfection. The curves for the 1:10 bleach, 1:100 bleach and sterile water controls are shown.

amount of bacteria recovered in the first twenty-four hours after treatment with the sodium hypochlorite. Treatment effectiveness was dependent on the strength of the treatment solution. The controls, with sterile water, also showed a decrease in the bacterial counts. The lack of nutrients in the sterile deionized water, the change in pH, or air purging the lines might account for the lowering of colony counts in the controls. Plot 2 displays the bacteria reductions over the course of the study. Both the 1:100 NaOCl and 1:100 NaOCl treated units resulted in at least a 2 log

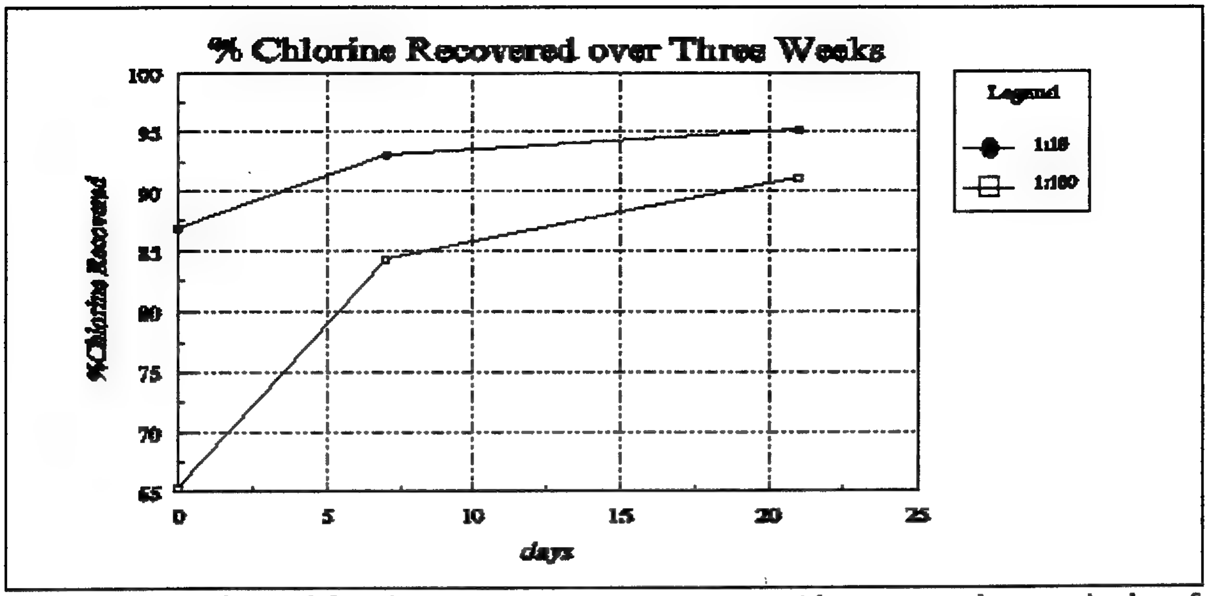


Plot 2: Averaged bacteria counts for the three sets of units over the course of three weeks.

reduction in colony counts over the twenty-one day period. In contrast, the control group showed approximately a 1.4 log reduction. The 1:10 has a quick kill while the 1:100 is more linear. The bacterial counts were also correlated with chlorine recovered after treatment. In contaminated lines, there was a

measurable decrease in the amount of chlorine recovered from units after each sodium hypochlorite treatment, which may be due to absorption of the chlorine by the bio-film adherent to the water lines. The greater the amount of bacteriological contamination present, as indicated by CFU/mL in pooled water samples, the greater degree of absorption. The chlorine consumption from the units at Brooks, Kelly, and Randolph was compared to those at Lackland. Brooks, Kelly and Randolph AFB's averaged 97%, 93%, and 80% chlorine recovery respectively. Randolph's low 80% recovery suggests that the units were highly contaminated[17], probably due to inadequate disinfection procedures. In comparison, Lackland's initial chlorine recovery averaged 86.9% for the 1:10 solution and 65.3% recovery for the 1:100[17]. Plot 3 illustrates the increases in chlorine recovery over the three weeks of cleaning. The higher recovery suggests that there is less bio-film to absorb the chlorine. This is consistent with Plot 2; as the colony counts approach zero, the chlorine recovery approaches 100%.[17] It appears that at the lower chlorine level 1:100, there is a higher percentage of the chlorine being absorbed; however, because there was a tenfold difference in the chlorine concentration between the two solutions, a greater amount of chlorine was absorbed from the 1:10 solution. It might be possible, after further study, to be able to estimate the amount of bio-film in the lines by the amount of chlorine recovered from cleaning.

All of the lines were studied under both light microscopy and



Plot 2: The chlorine recovery over a three week period of disinfection.

scanning electron microscopy. The LM was inconclusive in identifying bio-film in the lines; however, the analysis confirmed whether the lines contained deposits. The SEM, on the other hand, was very successful in identifying biological organisms and distinguishing bio-film from the inorganic deposits in the lines. Photo 1 shows one example of a contaminated line with bacteria growing on it. It also offers a perspective for the size relationship between biological and inorganic deposits on the lines. Many different types of organisms were found to be present, but it was not possible to identify specific bacteria. Photo 2 displays a variety of microbial morphologies, including spiral forms, cocci, and rods.

After two weeks of sodium hypochlorite treatment, there were no bacteria evident in the lines on SEM analysis, however, the

lines were not free of inorganic deposits. The ADEC units utilized two different types of materials in the lines that were studied: the inlet and air-water syringe lines were made with polyurethane without a plasticizer while the handpiece lines were polyvinyl chloride (PVC) with a high molecular weight plasticizer. After two weeks, the handpiece lines showed a significant

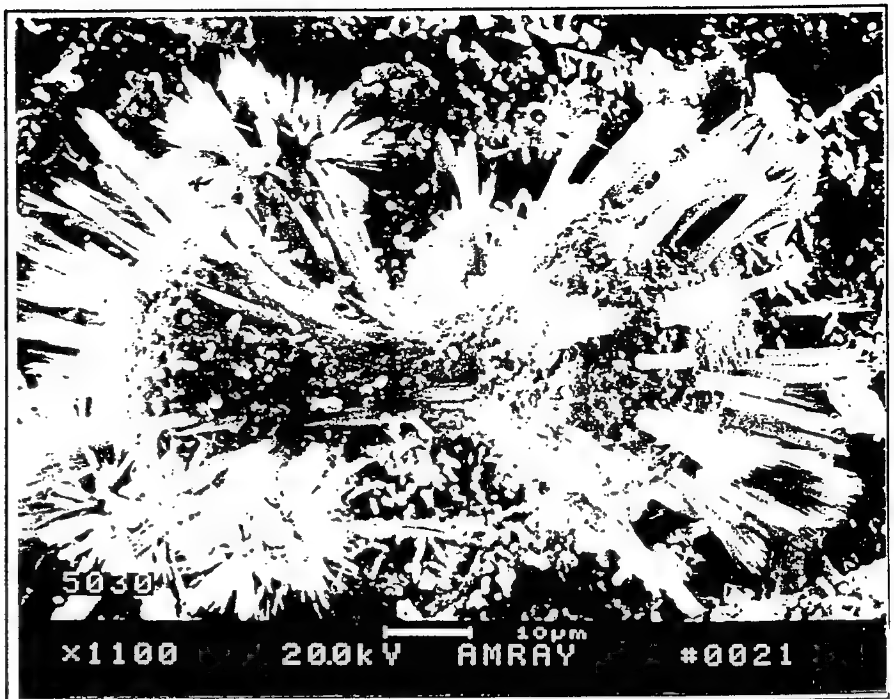


Photo 1: Example of an inorganic crystal with bacterial cells in the background and on the crystal itself.

Photo 2: This photo illustrates a variety of bacterial growth on the water line.

decrease in the deposits found on the lines. In contrast, the inlet and air-water syringe lines contained more visible deposits than before decontamination. The new deposits appear to be inorganic and may be the result of an initial reaction between the hypochlorite solution with either the water line itself, the bio-film, or other pre-existing mineral deposits.

The study was continued for four week time period of the study, the bacteria counts continually decreased in all of the units and negative results began to appear in the TC, LM, and SEM at three weeks. At the end of three weeks, the bacterial counts had dropped from an average of greater then 10^5 to 10^3 or less in non-control units. Controls also demonstrated a 1 log decrease.

Conclusions

All the dental units showed bio-film growth if they were not being treated consistently and accurately. This inconsistency could be eliminated by utilizing the correct procedure all of the time on all of the lines and units. Both 1:10 and 1:100 bleach solutions were effective in reducing bacterial contamination at Lackland AFB. This study suggests two possible methods for assessing the effectiveness of water line disinfection procedures: SPC Samplers and chlorine analyses. It was possible to determine if there was bio-film in the lines by determining the amount of chlorine that was recovered from the unit or by using an SPC sampler and performing the biological count.

Acknowledgements

The author would like to acknowledge and thank the Air Force Office of Scientific Research and Armstrong Laboratory for providing funding for the project. The author would also like to thank the Dental Investigation Service at Brooks AFB for their support. In addition, the author acknowledges her colleagues on the project: Drs. L.R. Sherman, S.E. Mills, and T.J. Plamondon.

References

1. Williams, J.F., et al. JADA. 124:59-65 (1993).
2. Reinhardt, R.A., R.W. Bolton, and G. Hlava. J. Periodont. 53: 96-100 (1982).
3. Michell, R, and M. Borneff. Zbl. Bakt. Hyg. 187:312-323 (1989)
4. Babcock, J. and M. Kleven. Unpublished work. Veterans Administration Clinic, Wood, WI, (1977).
5. Safe Drinking Water Act, Amended 1986; 42 USC. Public L.99-399.
6. Center for Disease Control and Prevention (CDC). Morbidity and Mortality Weekly Reports. 42: Nr. PR-8:7, (1993).
7. Fotos, P.G., et al. J. Hosp. Infec. 64:1382-1385 (1985).
8. Reinthaler, F.F., F. Mascher, and D. Stunzner. J. Dent. Res. 67:942-943 (1988).
9. Center for Disease Control (CDC). Morbidity and Mortality Weekly Reports. 39: Nr. 7:108-109 (1990).
10. Schied, R.C., et al. JADA. 105:658-660 (1982).
11. Dayoub, M.B., D.J. Rusilko, and A. Gross. J. Periodont. 49: 261-265 (1978).
12. Pankhurst, C.L., et al. J. Hosp. Inf. Cont. 16:9-18 (1990).

13. McEntegart, M.G. and A. Clark. Brit. Dent. J. 134:140-142 (1973).

14. Mills, S.E., P.W. Lauderdale, and R.B. Mayhew. JADA. 113: 280-284 (1978).

15. Moriarity, J.D. and J.J. Crawford. J. Dent. Res. 55: abstract nr. 855 (1976).

16. Sherman, L.R., S. Mills, and T. Plamondon, J. PA Acad. Sci. 69 iss. 1 (1995).

17. Sherman, L.R. Summer Research Project Final Report 33 (1995).

Appendix A:

	LA207	LA215	LC208	LC215	LA204	LA206	LC217	LC218	LA208	LA214	LC214	LC216
METAL	1:10	1:10	1:10	1:10	1/100	1/100	1/100	1/100	CON	CON	CON	CON
Cr-A	0	0	0	0	0	0	0	0	0	0	0	0
Cr-D	0	0	0	0	0	0	0	0	0	0	0	0
Cr-7d	0	0	0	0	0	0	0	0	0	0	0	0
Cr-14d	0	0	0.012	0	0	0	0	0	0	0	0	0
Cr-21d	0	0	0	0	0	0	0	0	0	0	0	0
Cu-A	0.318	0.562	0.419	0.317	1.83	0.41	0.251	0.263	0.374	0.449	0.335	0.393
Cu-D	1.05	1.07	0.858	0.47	1.84	0.82	1.07	0.324	1.26	0.891	0.641	0.885
Cu-7d	0.163	0.074	0.107	0.049	0.097	0.038	0.142	0.028	0.026	0.056	0.048	0.076
Cu-14d	0.415	0.126	0.158	0.243	0.124	0.092	0.287	0.192	0.081	0.081	0.157	0.045
Cu-21d	0.216	0.224	0.063	0.073	0.076	0.033	0.106	0.075	0.062	0.063	0.137	0.07
Ni-A	0.584	0.075	0.061	0.63	0.146	0	0.015	0.123	0	0.094	0.076	0.054
Ni-D	0.238	0.06	0.427	0.138	0.12	0.039	0.155	0.078	0.286	0.127	0.038	0.23
Ni-7d	0	0	0	0	0.036	0	0	0	0	0	0	0
Ni-14d	0.062	0	0.052	0.033	0	0	0.058	0.056	0	0	0	0
Ni-21d	0.048	0	0.035	0.035	0	0	0.038	0	0	0.039	0.036	0
Zn-A	0	0.08	0.066	0.11	1.45	0.056	0.06	0.136	0	0.059	0.12	0.089
Zn-D	0.084	0.185	1.13	0.231	1.49	0.151	0.352	0.181	0.163	0.483	0.109	0.359
Zn-7d	0.106	0.074	0.167	0.075	0.123	0	0.095	0	0	0.094	0.106	0.127
Zn-14d	0.563	0.157	0.194	0.237	0.27	0.081	0.218	0.204	0.28	0.315	0.229	0.083
Zn-21d	0.244	0.184	0.132	0.125	0.314	0.082	0.793	0.155	0.103	0.262	0.204	0.089
Ca-A	47.4	76.6	74.5	77.2	55.7	75.7	85.3	79.1	75.4	76.1	73.5	75.8
Ca-D	53.2	54.7	54.7	51.4	85.5	50.8	79.4	50	55.4	58.2	50	63.6
Ca-7d	0	0	0	0	0	0	0	0	0	0	0	0
Ca-14d	0	0	0	0	0	0	0	0	0	0	0	0
Ca-21d	0	2.26	1.78	1.5	0	0	3	1.64	0	0.33	0	1.37

Appendix B

SAMPLE #	TYPE	LM	TC	SEM	AO	SPC	L
		I/A2/H	I/A2/H	I/A2/H	I/A2/H		
B153	1\10	-/+/+	-/-/-	-/-/-	NO SAM.	-: 0	NO SAM.
B155	1\10	+/-/+	-/-/-	?/?/+	NO SAM	-: 0	NO SAM
B158	1\10	+/+/+	-/-/-	+/+/-	NO SAM.	+: 30	NO SAM.
B166	1\10	-/-/+	-/-/-	-/+/+	NO SAM.	+: 30	NO SAM.
K124	1\10	-/-/+	-/-/+	+/+/+	NO SAM.	+: 8820	NO SAM.
K125	1\10	/ /	-/+/+	/ /	NO SAM.	-: 0	NO SAM.
K128	1\10	-/+/+	-/-/+	-/+/-	NO SAM.	-: 0	NO SAM.
K129	1\10	-/+/+	+/ /+	-/+/?	NO SAM.	-: 0	NO SAM.
R118	1\10	+/+/+	+/+/-	/ /	NO SAM.	+: 110	NO SAM.
R212	1\10	+/+/-	+/+/-	/ /	NO SAM.	+: 47880	NO SAM.
R214	1\10	+/+/-	+/+/-	/ /	NO SAM.	+: 18480	NO SAM.
R219	1\10	+/+/-	+/+/+	/ /	NO SAM.	+: 5540	NO SAM.

Appendix C

BACTERIA

ROOMS	DECON	D+1D	D+2D	D+5D	D+7D	D+14D	D+21D
LA207:10	757000	150	220	999	2000	190	2490
LA215:10	56000	80	140	1000	6000	220	0
LC208:10	301000	90	20	5000	7000	2850	0
LC215:10	234000	10	100	1000	1000	20	10
LA204:100	124000	840	32110	60000	22000	4800	1400
LA206:100	138000	690	7200	27000	53000	10200	1600
LC217:100	185000	430	170	6000	71000	30800	3000
LC218:100	260000	890	13800	3000	35000	19000	0
LA208:C	117000	37650	168000	145000	61000	13000	18000
LA214:C	1061000	5440	83160	53000	43000	39000	32000
LC214:C	106000	12600	84000	47000	65000	24000	0
LC216:C	234000	29000	98280	211000	348000	15000	2000

Appendix D

SAMPLE	TYPE	LM	TC	SEM	AO	SPC	L
		I/A2/H	I/A2/H	I/A2/H	I/A2/H		
LA207-6/20	1\10	+/+/+	?/+/--	+/ /	+/+/+	+	NO SAM.
LA207-7/5	1\10	+/+/+	+/+/+		NO SAM.	+: 757000	NO SAM.
LA207-7/12	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 2000	-
LA207-7/19	1\10	+/+/+	+/-/-			+: 190	NO SAM.
LA207-7/26	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 2490	
LA215-6/20	1\10	+/+/+	?/?/+		+/+/+	+	NO SAM.
LA215-7/5	1\10	+/+/+	+/+/+		NO SAM.	+: 56000	NO SAM.
LA215-7/12	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 7000	-
LA215-7/19	1\10	+/+/+	√/-/-			+: 2850	NO SAM.
LA215-7/26	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	-: 0	
LC208-6/20	1\10	+/+/+	--/+/--		+/+/+	+	NO SAM.
LC208-7/5	1\10	+/+/+	+/+/+		NO SAM.	+: 301000	NO SAM.
LC208-7/12	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 7000	-
LC208-7/19	1\10	+/+/+	+/+/+			+: 2850	NO SAM.
LC208-7/26	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	-: 0*	
LC215-6/20	1\10	+/+/+	?/+/+	/ /?	+/+/+	+	NO SAM.
LC215-7/5	1\10	+/+/+	+/+/+		NO SAM.	+: 234000	NO SAM.
LC215-7/12	1\10	NO SAM.	NO SAM	NO SAM.	NO SAM.	+: 1000	-
LC215-7/19	1\10	+/+/+	+/+/+			+: 20	NO SAM.
LC215-7/26	1\10	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 10*	
LA204-6/20	1\100	+/+/+	+/?/?	+/+/?	+/+/+	+	NO SAM.
LA204-7/5	1\100	+/+/+	+/+/+		NO SAM.	+: 124000	NO SAM.
LA204-7/12	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 22000	+
LA204-7/19	1\100	+/+/-	-/+/-	-/-/+		+: 2400	NO SAM.
LA204-7/26	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 700	
LA206-6/20	1\100	+/+/+	?/?/?	/ /+	+/+/+	+	NO SAM.
LA206-7/5	1\100	+/+/+	+/+/+		NO SAM.	+: 138000	NO SAM.
LA206-7/12	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 53000	-
LA206-7/19	1\100	+/+/+	+/+/+	-/-/?		+: 51000	NO SAM.
LA206-7/26	1\100	NO SAM.	NO SAM.		NO SAM.	+: 800	
LC217-6/20	1\100	+/+/+	+/?/+	+/+/+	-/-/-	+	NO SAM.
LC217-7/5	1\100	+/+/+	+/+/+		NO SAM.	+: 185000	NO SAM.
LC217-7/12	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 71000	-
LC217-7/19	1\100	+/+/+	√/-/-			+: 15400	NO SAM.
LC217-7/26	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 1500*	
LC218-6/20	1\100	+/+/+	+/?/+		+/+/+	+	NO SAM.
LC218-7/5	1\100	+/+/+	+/+/+	/ /+	NO SAM.	+: 260000	NO SAM.
LC218-7/12	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 35000	-
LC218-7/19	1\100	+/+/-	+/+/+			+: 9500	NO SAM.
LC218-7/26	1\100	NO SAM.	NO SAM.	NO SAM.	NO SAM.	-: 0*	

SAMPLE	TYPE	LM	TC	SEM	AO	SPC	L
		I/A2/H	I/A2/H	I/A2/H	I/A2/H		
LA208-6/20	CONTROL	+/+/+	–/+/–	+/?/+	+/-/+	+	NO SAM.
LA208-7/5	CONTROL	+/+/-	+/+/+		NO SAM.	+: 117000	NO SAM.
LA208-7/12	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 61000	-
LA208-7/19	CONTROL	+/+/+	+/+/-			+: 13000	NO SAM.
LA208-7/26	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 18000	
LA214-6/20	CONTROL	+/+/+	+/–/–	+/+/+	+/+/+	+	NO SAM.
LA214-7/5	CONTROL	+/+/+	+/+/+		NO SAM.	+:1061000	NO SAM.
LA214-7/12	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 43000	-
LA214-7/19	CONTROL	+/+/+	+/+/+			+: 39000	NO SAM.
LA214-7/26	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 32000	
LC214-6/20	CONTROL	+/+/+	–/–/+		+/+/+	+	NO SAM.
LC214-7/5	CONTROL	+/+/+	+/+/-		NO SAM.	+: 106000	NO SAM.
LC214-7/12	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 65000	-
LC214-7/19	CONTROL	+/+/+	+/+/+			+: 24000	NO SAM.
LC214-7/26	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	-: 0*	
LC216-6/20	CONTROL	+/+/+	+/–/–		+/+/+	+	NO SAM.
LC216-7/5	CONTROL	+/+/+	+/+/+		NO SAM.	+: 234000	NO SAM.
LC216-7/12	CONTROL	NO SAM.	NO SAM.	NO SAM.		+: 348000	-
LC216-7/19	CONTROL	+/+/+	+/+/-			+: 15000	NO SAM.
LC216-7/26	CONTROL	NO SAM.	NO SAM.	NO SAM.	NO SAM.	+: 2000*	

* THESE UNITS RECIEVED AN EXTRA CLEANING OF 1:10 SODIUM HYPOCHLORITE TWO DAYS AFTER THE REGULAR CLEANING.

NOTE 1: ALL BLANK SPACES SIGNIFY DATA THAT WAS NOT COMPLETED AS OF 8/4/95.

NOTE 2: QUESTION MARKS SIGNIFY DATA THAT WAS INCONCLUSIVE AS OF 8/4/95.

METHODS FOR ESTABLISHING DESIGN LIMITS TO ENSURE ACCOMMODATION

Kristie J. Nemeth
Graduate Student
Department of Psychology

Miami University
128 Benton Hall
Oxford, Ohio 45056

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1995

METHODS FOR ESTABLISHING DESIGN LIMITS TO ENSURE ACCOMMODATION

Kristie J. Nemeth
Department of Psychology
Miami University

Abstract

Whenever a workspace is designed, one of the objectives should be to accommodate the user population. Typically, designers attempt to allow 90% of the population to reach controls, fit through spaces, and lift objects. This paper outlines potential methods for determining the chosen range of accommodation. After the traditional procedures are described and used to establish a design solution, an alternative method is recommended. This new multivariate technique requires validation testing prior to its implementation in the Design and Evaluation of Personnel, Training and Human Factors (DEPTH) software which is being designed at Wright Patterson Air Force Base, Dayton, Ohio..

METHODS FOR ESTABLISHING DESIGN LIMITS
TO ENSURE ACCOMMODATION

Kristie J. Nemeth

Review of Traditional Methods

Whenever a workspace is designed, one of the objectives should be to accommodate the user population. Typically, designers attempt to allow 90% of the population to reach controls, fit through spaces, and lift objects. The traditional test for accommodation is to ensure that the 5th percentile person be able to reach and lift everything and that the 95th percentile person fit everywhere. Unfortunately, in most cases this test methodology does not lead to 90% accommodation.

For simple design problems, using percentile extremes to test for accommodation may be appropriate. To test that 95% of the population can reach a control switch located overhead, it is appropriate to use the 5th percentile "overhead reach" as the upper limit of switch placement. Similarly, when developing the lowest doorway height that 95% of the population can fit through without stooping, it is appropriate to use the 95th percentile stature. If a person with 95th percentile stature can fit through the doorway, then anyone shorter will also fit.

Unfortunately, most design problems are more complicated. They require the use of multiple body dimensions for design specification development. To assess more complicated workspaces, it is often easier to generate human models. The models can then be used to evaluate specific tasks in the workspace. It has been unclear how multiple body dimensions should be combined to create human models. How is a 95th percentile person defined? One suggested method is to set each of the body segments to the 95th percentile value. A second method is to set the stature of the model to the 95th percentile and use regression equations to predict the size of each of the body segments. Each of these methods has its associated problems (for a review see Robinette & McConville, 1981). Both methods result in models

that do not accurately reflect human size or variability. They can also lead to inaccurate testing of accommodation.

<u>Design Example</u>

An examination of the design criteria for the location of control buttons on a vertical panel demonstrates the potential for error (see Figure 1). The task requires a 45 cm deep shelf that can be located at any height. To facilitate the analysis, assume that the person has to stand upright, and 15 cm from the shelf. The relevant dimensions for the task are standing shoulder height and thumbtip reach.

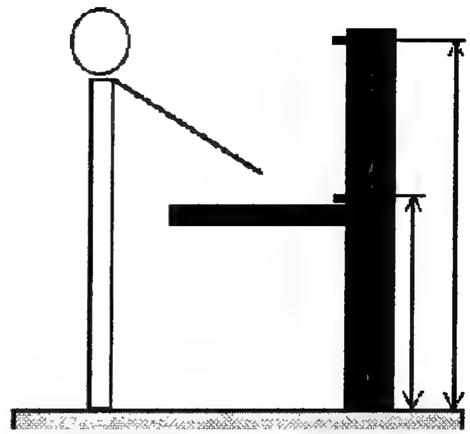


Figure 1. Example design problem: Locate upper and lower limits for button location.

Using the goal of 90% accommodation, traditional methods use 5th and 95th percentile models. However, both of the methods described earlier generate differently sized 5th and 95th percentile models. Table 1 lists the measurements and formulas used to determine reach ranges for both model generation methods. The dimensions are derived

<u>Table 1. Reach Range Calculations</u>

Dimension Measurements

	Standing Shoulder Height		Thumbtip Reach Length	
	5th Percentile	95th Percentile	5th Percentile	95th Percentile
Segment-Based	131.85 cm	156.65 cm	72.25 cm	87.92 cm
Stature- Based	134.72 cm [1]	152.68 cm	72.46 cm [2]	82.13 cm

Calculated Reach Limits

	Highest Reach		Lowest Reach	
	5th Percentile	95th Percentile	5th Percentile	95th Percentile
Segment-Based	172.10 [3]	220.91	91.60 [4]	92.39
Stature- Based	96.60 cm	208.76 cm	94.09 cm	175.35 cm

[1] Standing Shoulder Height = .818 * Stature (Drillis & Contini, 1966)
[2] Thumbtip Reach Length = .44 * Stature
[3] Highest Reach = $131.85 + (72.25^2 - (45+15)^2)^{1/2} = 172.10$
[4] Lowest Reach = $131.85 - (72.25^2 - (45+15)^2)^{1/2} = 91.60$

from the ANSUR (1988) anthropometric survey of Army males.

Figure 2 represents reach ranges for 5th and 95th segment-based percentile models. The optimal design range for 90% accommodation appears to fall between the 5th percentile upper reach limit and the 95th percentile lower reach limit. Considering this, the upper and lower limits for control placement will be based on this range. To test this assumption, the same formulas found in Table 1 were used to calculate the actual reach range for each male in the ANSUR (1988) anthropometric survey.

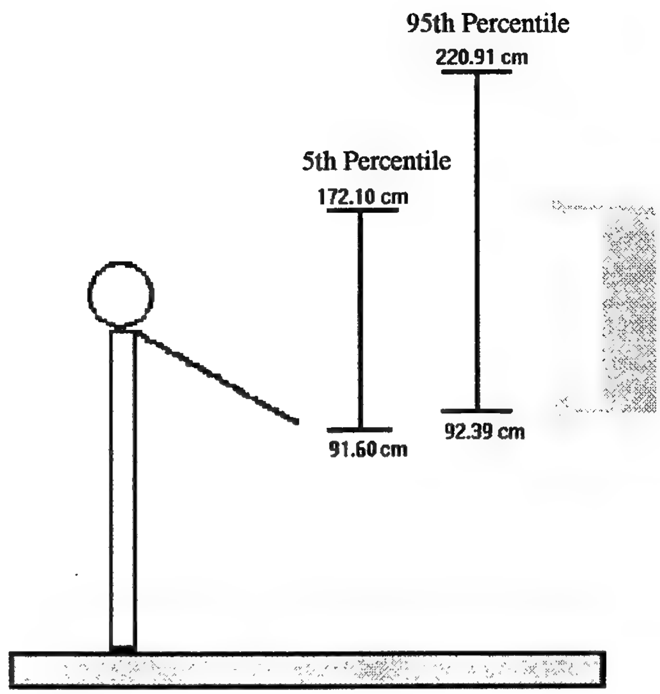


Figure 2. Reach ranges for models generated by segment-based Method.

Results of Percentile Methods

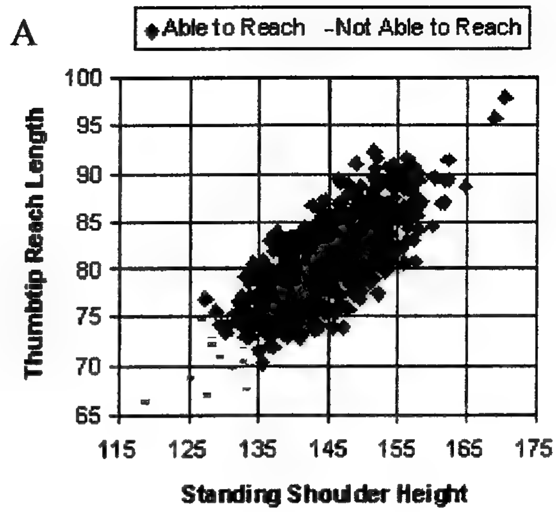
Figure 3 shows who could and could not actually reach the controls. Each point on the multivariate distribution represents one Army male in the ANSUR (1988) anthropometric survey. For example, subject #17689 has a standing shoulder height of 170.4 cm and a thumbtip reach length of 98 cm. He is able to reach the upper control, but cannot reach the lower control.

The goal of 90% accommodation requires that only 5% of the population be unable to reach the lowest control, and only 5% of the population be unable to reach the highest control. If these conditions are exceeded, space and resources are wasted.

The 5th percentile segment-based model has an upper reach limit of 172.10 cm. If we place the highest control button at this point, 98.53% of the population can reach it (Figure 3, A). This conservative upper reach limit accommodates a larger percentage of the population than called for by the specifications, and wastes 5 cm of potential workspace.

The 95th percentile segment-based model has a lower reach limit of 92.39 cm. If we place the lowest control at this point, only 59.75% of the population can reach it (Figure 3, B). Clearly, that falls short of the accommodation goal.

Inaccurate results also occur when generating models with the stature-scaling method. If we use the 5th percentile model to place the highest control buttons, and the 95th percentile model to place the lowest control buttons, we will accommodate 89.91% and 97.29%, respectively.



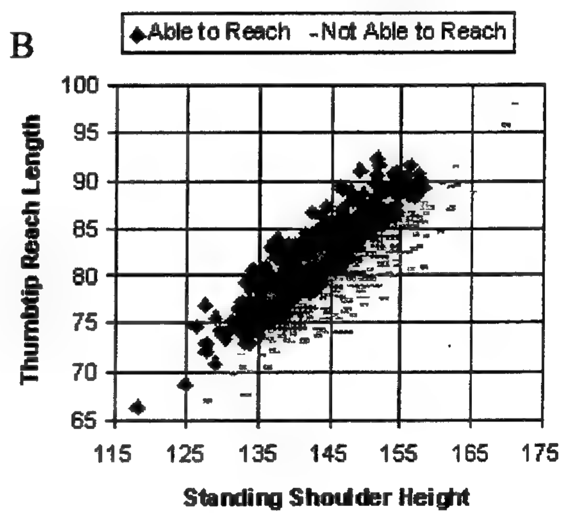


Figure 3. Results of design limits for upper (A) and lower (B) controls. Design range determined by Segment-Based model reach ranges.

Individual Data instead of Percentiles

Obviously a different technique is necessary to accurately determine design specifications for a given range of accommodation. Using the reach range computed for each subject in the ANSUR (1988) anthropometric survey of Army males, design specifications can be obtained which come much closer to reaching the accommodation goal of 90%. If the highest control buttons are placed at 178.50 cm, only 5% of the sample will be unable to reach them. If the lowest control buttons are placed at 98 cm, only 5% of the sample will be unable to reach them. Assuming that the people who cannot reach the high control buttons are not the same people who cannot reach the low control buttons, 90% of the sample population is accommodated by locating the control buttons between 98 and 178.50 cm.

Multivariate Method

Using the raw data in this manner to test each subject's reach, fit, and strength could become overwhelming and time consuming for the typical designer. Furthermore, accommodation tests that use more than two relevant dimensions cannot be solved with simple geometry. To overcome these problems, a multivariate technique has been recommended (Zehner, Meindl, & Hudson, 1993).

The first step is to identify the target population and the appropriate anthropometric survey. The second step is to establish accommodation goals. In the example above, the goal was to accommodate 90% of the males from the ANSUR (1988) anthropometric survey.

The next step, identifying the relevant dimensions, depends on the task to be evaluated. As an example, to determine seat height ranges, popliteal height is important. But popliteal height is unimportant for back rest height ranges. Most design problems have several relevant dimensions. A Principle Component Analysis can determine the two (or three) components that account for most of the variance in the relevant dimensions. It is

then possible to generate a multivariate distribution from these two (or three) factors. This statistical analysis will convert an elliptical distribution to a circular (or spherical) distribution, as shown in Figure 4.

When this distribution is defined, it becomes possible to determine an accommodation circle. The circle in Figure 4 encompasses 90% of the distribution. 90% of the sample members fall inside the circle and 10% fall outside. Reaching the accommodation goal of 90% requires accommodating everyone who falls inside the circle.

If we select representative cases from the perimeter of the circle, it is possible to determine whether the models are able to reach, fit or lift as necessary to perform the task and make design recommendations as necessary (see Figure 5). Sampling throughout the distribution is also recommended.

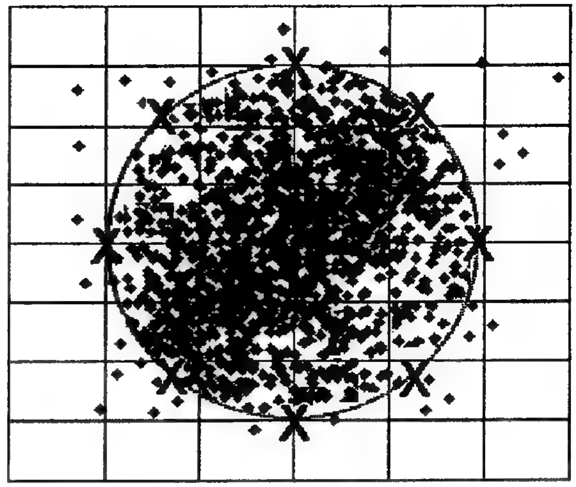


Figure 4. Example distribution of principle components for a specific design task. Eight representative cases from perimeter of accommodation circle are marked.

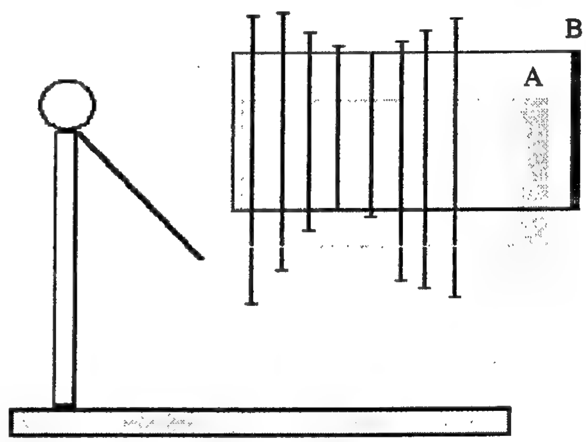


Figure 5. Reach ranges of eight representative models from the perimeter of the accommodation circle are shown. The design range derived from the percentile models is inadequate (A). The design range selected from the eight reach ranges shown accurately accommodates 90% of the population (B).

<u>Conclusions</u>

By generating computer models of humans with dimensions related to

the representative cases, we create a collection of models to use in the task evaluation. Instead of using two fabricated models (5th and 95th percentile models), models conforming to the dimensions of specific, real individuals are used to perform the task. Using those representative models, the new method reflects the variability of humans, and can be used as a more accurate test for accommodation in a workspace.

References

Drillis, R., and Contini, R. (1966). Body Segment Parameters, BP174-945, Technical Report TR 1166.03, School of Engineering and Science, New York University, New York.

Gordon, C.C., Bradtmiller, B., Clauser, C.E., Churchill, T., McConville, J.T., Tebbetts, I., and Walker, R.A. (1989). 1987-1989 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics. Technical Report NATICK/TR-89-044, U.S. Army Natick Research, Development and Engineering Center, Natick, Massachusetts.

Robinette, K.M., and McConville, J.T. (1981). An Alternative to Percentile Models. Society of Automotive Engineers Technical Paper Series, International Congress and Exposition. Detroit, Michigan.

Zehner, G.F., Meindl, R.S., and Hudson, J.A. (1993). A Multivariate Anthropometric Method for Crew Station Design: Abridged (U). Technical Report AL-TR-1992-0164. Crew Systems Directorate, Human Engineering Division, Armstrong Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio.

EVALUATION OF A NEW DERIVATIZING REAGENT, 1-(9-ANTHRACENYLMETHYL)PIPERAZINE (MAP), USED FOR THE ANALYSIS OF ISOCYANATES IN SPRAY-PAINTING OPERATIONS

Samuel H. Norman
Graduate Student
Department of Chemistry

Southwest Texas State University
San Marcos, Texas 78666

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1995

EVALUATION OF A NEW DERIVATIZING REAGENT, 1-(9-ANTHRACENYLMETHYL)PIPERAZINE (MAP), USED FOR THE ANALYSIS OF ISOCYANATES IN SPRAY-PAINTING OPERATIONS

Samuel H. Norman
Graduate Student
Department of Chemistry
Southwest Texas State University

Abstract

NIOSH Method 5521 and a method currently being developed by NIOSH were used for the sampling and analysis of 1,6-hexamethylene diisocyanate, or HDI, monomer and oligomer. Field studies were conducted during several spray-painting operations, and NIOSH Method 5521 which uses 1-(2-methoxyphenyl)piperazine, or MOP, as a derivatizing reagent was compared directly with the new NIOSH method which employs 1-(9-anthracenylmethyl)piperazine, or MAP. Spray painting operations were conducted at Keesler Air Force Base in Mississippi. HDI monomer concentrations for these operations were all below the limit of detection for both the MOP and the MAP samples. HDI oligomer, or HDI-based polyisocyanate, was detected easily for all samples using either derivatizing reagent. The average polyisocyanate concentration in air as measured by the MAP-polyisocyanate urea response varied between 1.1 and 6.1 mg/m^3. The polyisocyanate concentration as measured by the MOP-polyisocyanate urea response varied between 1.4 and 4.6 mg/m^3. To compare spatial variation for the MAP derivatizing reagent, two side-by-side MAP/DMSO impingers were placed within five inches of each other during sampling. For the 12 data pairs in this paper, the average deviation from each other is 0.38 mg/m^3 (n=12). When comparing temporal variation, the pooled average deviation of a sample when run over a ten day period is 0.51 mg/m^3 (n=14). When comparing the two MAP results with those obtained from the MOP analysis, the results appear to be comparable.

EVALUATION OF A NEW DERIVATIZING REAGENT, 1-(9-ANTHRACENYLMETHYL)PIPERAZINE (MAP), USED FOR THE ANALYSIS OF AIRBORNE ISOCYANATES IN SPRAY-PAINTING OPERATIONS

Samuel H. Norman

Introduction

Isocyanates are major components in polyurethane paint formulations because of their ability to form durable crosslinks in the polyurethane coating. The most popular isocyanate employed for this hardening ability is 1,6-hexamethylene diisocyanate (HDI) both in its monomeric and oligomeric forms. The monomeric form of HDI was used very much in the past, however due to its toxicity, HDI monomer has been highly regulated by the American Council of Governmental Industrial Hygienists (ACGIH).[1] Because regulations were imposed on the use of HDI monomer as a paint hardener, industries started using the HDI oligomer (mainly biuret or isocyanurate of HDI) as the primary hardener in paint formulations. The oligomers were thought to be less harmful because they were less volatile than the monomer and thus less reactive in the human body. However, recent reports have indicated that HDI oligomer, or polyisocyanates, can cause occupational asthma as well as other respiratory problems for workers involved in spray-painting operations.[2-4]

Because of these health risks involved with use of polyisocyanates in spray paint formulations, routine chemical analysis of the concentration level of airborne polyisocyanates during spray-painting operations will soon be needed. The most important factor in an isocyanate chemical analysis is the choice and use of a suitable derivatizing reagent. A suitable derivatizing reagent for isocyanate analysis would possess the following characteristics: 1) High reaction rate with isocyanate during analysis, 2) The derivatized isocyanate must be UV sensitve at low concentrations, 3) The derivatized isocyanate must be selective by separating it from interferants, and 4) The derivatized isocyanate must be detectable outside of the isocyanate structure. Several different derivatizing reagents for the derivatization of isocyanates have been used and are currently being used in isocyanate analysis. While most of these reagents work well to derivatize isocyanates and offer good sensitivity and selectivity in the isocyanate sampling and analysis[5-6], the search for a better and more efficient derivatizing reagent will continue. 1-(9-anthracenylmethyl)piperazine (MAP) is a novel derivatizing reagent that possesses all of the above characteristics and appears to be a much better reagent than any of the reagents before it. All comparisons in this report are made between MAP and the derivatizing reagent currently being used for isocyanate analysis

1-(2-methoxyphenyl)piperazine (MOP). When compared to MOP, MAP's qualities include the following: 1) A similar yet higher reaction rate, 2) 30% more sensitive to UV detection, 3) Added fluorescence detection for greater selectivity, and 4) greater distance between isocyanate group attachment and chromophore/fluorophore to establish independence of isocyanate structure and derivatizing reagent.(6) Although MAP appears to be a much better isocyanate derivatizing reagent, testing and method development has not been completed. The scope of the research in this report is to compare the performance and efficiency of MAP with that of the current reagent, MOP.

In the following report, field samples were obtained during several spray-painting operations for HDI monomer and oligomer analysis. The performance and efficiency of MAP was compared to that of MOP by placing one MOP sampling device and one MAP sampling device next to each other during sampling events. A third sampling device containing MAP was placed next to the other two in an effort to study the MAP sampling variations. Finally, an indirect comparison of methods and method development is included in this report.

Experimental

Methods

Several Methods were used to determine the amount of isocyanates generated during spray-painting operations. NIOSH Method 5521 and a new method currently being developed by NIOSH were used for the sampling and analysis of airborne HDI monomer and oligomer. NIOSH Method 5521, modified so that polyisocyanate concentrations were determined based on polyisocyanate standards, was used to determine the concentration of polyisocyanate.

Reagents and Apparatus

1-(2-methoxyphenyl)piperazine (MOP) was obtained from Fluka; and 1-(9-anthracenylmethyl)piperazine (MAP) was synthesized and purified according to the procedure given by NIOSH.(1) Desmodur N-75, which is 75% HDI polyisocyanate in xylene and contains 35-40% biuret of HDI (MSDS), was obtained from Miles Inc. (Pittsburgh, PA). Bulk catalysts were Deft polyurethane paint catalyst which is 60% HDI polyisocyanate in xylene. Bulk catalyst samples were used to prepare polyisocyanate standards and were sent with the corresponding field samples. The 98% pure HDI monomer was purchased from Kodak Company. All other chemicals and solvents were reagent grade or better.

Standard midget impingers filled with solutions of 43 mg/L of MOP in toluene or 62 mg/L of MAP in DMSO were used to collect isocyanate.(14) Dupont Alpha 1 pumps were used to draw

the air samples. The pumps were calibrated before and after sampling with a Gillian bubble generator. A high pressure spray gun was used for operation 1K. High volume low pressure (HVLP) spray guns were used in operations 2K and 3K.

Description of Spray Paint Operations

All spray-painting operations were batch, performed by different painters who used supplied air hoods to cover the head and TIVEC suits to cover the body. Three different spray paint operations were evaluated at Keesler AFB, MS. Operation 1K involved painting a van type vehicle blue in a drive-in style spray paint booth. Operation 2K involved spray-painting a C-130 aircraft engine gray in a three-sided spray paint area with a waterfall type air cleansing system at the front of the bay. Operation 3K involved spray-painting a power generator green in the same type of spray paint area as Operation 2K.

Area sampling was performed by positioning two impingers containing MAP in DMSO next to a third impinger containing MOP in toluene. The inlets were located about 4.0-4.5 in. apart from each other. The samples at Keesler AFB were collected in an area in which the painter routinely does not work, but in an area of overspray. Area samplers were positioned 4.0-4.5 ft. above the floor, and about 5-6 ft. downdraft from the equipment being painted. For painting areas which used a waterfall type air cleansing system, area samplers were positioned in front of the waterfall.

The sampling protocols described in NIOSH Method 5521 were used throughout. The isocyanate samples were collected at a flow rate of 1.0 L/min., and sampling times ranged from 10-63 minutes.

Analytical Instrumentation

The HPLC system consisted of a Hewlett-Packard Series 1090 chromatograph with autosampler and diode array UV-VIS detector operated at 246 nm for MOP and 254 nm for MAP. A Hewlett-Packard 1049A electrochemical detector operated in the amperometric mode at +0.8 V was connected in series with the HPLC. Hewlett-Packard 3396 series II integrators were used to determine the area under all chromatographic peaks.

A Hewlett-Packard LiChrosorb RP-18, 10 μm (2200 X 4.6 mm) column and (45/55) acetonitrile/methanolic buffer (pH=6.0) mobile phase as specified in NIOSH Method 5521 were used for analyzing the MOP derivatives. A Hewlett-Packard LiChrosorb RP-8, 10 μm (2200 X 4.6 mm) column and (58/42) acetonitrile/0.04 M sodium phosphate buffer adjusted to a pH of 2.85 with phosphoric acid were used for analyzing the MAP derivatives. In order to speed up the

analysis of the MAP derivatives, the pH was adjusted to 2.50. The mobile phase flow rate for the HDI polyisocyanate derivatives was 2.000 mL/min. The flow rate for HDI monomer was 1.000 mL/min. The sample injection volume was consistently 20 μL.

Preparation of Standards

MOP-HDI standards were prepared as described in NIOSH Method 5521. MAP-HDI standards were prepared as described by NIOSH.[6]

MOP-polyisocyanate and MAP-polyisocyanate standards for method development work were prepared from the bulk polyisocyanate-based catalysts (Deft hardener) that were used in the spray-painting operations. Separate catalyst samples were sent with each batch. The bulk catalyst sample from each batch was blown down until further evaporation was not observed. Stock solutions of 1000 ppm were prepared as follows: 50 mg of derivatizing reagent was dissolved in 50 mL of toluene, and 25 mg of prepared bulk catalyst was dissolved in 50 mL of toluene. The two solutions were combined and allowed to react overnight or until the solution turned clear. The excess derivatizing reagent was acetylated with 70 μL of acetic anhydride, and then the toluene was evaporated at 60 °C under a gentle stream of nitrogen. The polyisocyanate derivatives were reconstituted in 25 mL of methanol for the MOP derivatives and 25 mL of DMSO for the MAP derivatives.

After collection, isocyanate samples in MOP/toluene were immediately refrigerated at 4 °C, whereas the samples in MAP/DMSO were maintained at room temperature. Both sample types were then shipped in screw-cap vials within two days of sampling. Upon arrival at the analytical lab, the samples were refrigerated, and then prepared within three days. The samples were analyzed within 2-5 weeks and then repeatedly within a two week period.

Analysis

In order to compare the results obtained for HDI monomer and HDI-based polyisocyanate using the two derivatizing reagents MOP and MAP, both were quantitated using the UV detector response. In one situation, however, the electrochemical detector (ECD) response was used because both an interferant was present on the UV response, and the resulting calibration curve was much more linear. The ECD response was also used to confirm the presence of isocyanates or moieties that contain isocyanate groups.

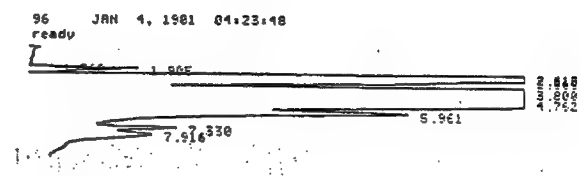
The concentration of polyisocyanate in field samples was determined by using the area under the peak in the chromatogram of a sample and comparing it with the area under the corresponding peak in the chromatogram of a bulk catalyst standard. Calibration curves based on

MOP-HDI urea, MAP-HDI urea, MOP-polyisocyanate, and MAP-polyisocyanate urea standards were all linear with a correlation coefficient of r > 0.9975.

Results and Discussion

HDI monomer concentrations during spray paint operations at Keesler AFB in Mississippi were all below the limit of detection whether using MOP or MAP as the derivatizing reagent. Although MAP has a molar absorptivity which is thirty times as high as that of MOP, the sensitivity advantage associated with MAP was obviated because of the large tailing associated with the acetylated MAP derivatizing reagent (See Figure 1). Repeated efforts to extract the HDI-MAP urea peak from the acetylated MAP tail were ineffective.

Figure 1. Large tailing associated with the acetylated form of MAP which prevented separation and detection of HDI monomer {Sample # 276MAP(63L)}. Run conditions for are as reported in the Experimental Section.



HDI-based polyisocyanate concentrations were easily accessible with either the MAP or MOP reagents. Good chromatographic separation and resolution was achieved with both the MOP and MAP polyisocyanate analyses. The chromatograms in Figure 2 show the good separation of the MOP and MOP polyisocyanate ureas from the acetylated derivatizing reagent and other interferants in the sample.

Figure 2. Separation of a)MOP- and b)MAP- polyisocyanate ureas from acetylated peak and other interferants within the field sample {Sample # 276MAP(63L)}. Run conditions are as reported in the Experimental Section. Derivatized polyisocyanate is denoted with an asterick*.

a)

b)

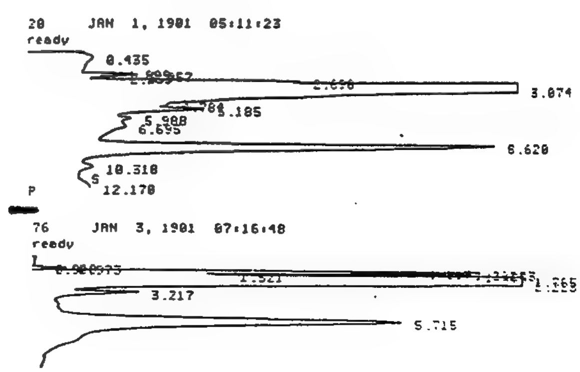


Table I reports the polyisocyanate air concentration obtained from six different sampling events during three different spray paint operations.

TABLE I. Airborne Polyisocyanate Concentrations in the Spray-Painting Environment (Air Concentration in mg/m^3)

Operation (Date)	Sample Number	Run #1	Run #2	Run #3	Run #4	Average
1K (6-21-95)	#273MAP(62L)	6.5	6.7	6.1	5.0	**6.0**
	#274MAP(60L)	6.3	6.8	6.0	5.4	**6.1**
	#275MOP(11L)					**4.6**
1K (6-21-95)	#276MAP(63L)	5.0	5.1		4.8	**5.0**
	#277MAP(61L)	4.6	4.7		4.9	**4.7**
	#278MOP(14L)					**4.4**
2K (6-27-95)	#156MAP(21L)	3.7	2.2			**3.0**
	#157MAP(20L)	4.6	2.8			**3.7**
	#158MOP(23L)					**3.6**
2K (6-27-95)	#159MAP(42L)	1.9	1.2			**1.5**
	#160MAP(42L)	2.0	1.4			**1.7**
	#161MOP(47L)					**1.4**
3K (7-11-95)	#527MAP(39L)	---	1.2	1.0	---	**1.1**
	#528MOP(39L)					**1.7**
3K (7-11-95)	#531MAP(11L)	5.9	---	---	6.2	**5.9**
	#529MAP(10L)	7.0	4.6	4.3	—	**5.4**
	#530MOP(11L)					**3.1**

The average polyisocyanate concentration in air as measured by the polyisocyanate urea response varied between 1.1 and 6.1 mg/m^3 (see column under MAP). The polyisocyanate concentration as measured by the polyisocyanate-MOP urea response varied between 1.4 and 4.6 mg/m^3 (see column under MOP). All of the values exceed 1 mg/m^3 which is the short-term exposure limit (STEL) recommended by the manufacturer (Bayer) and the ceiling standard established by OSHA in Oregon.[7] At the present time there are no federal OSHA permissible exposure limits (PEL) or short-term exposure limits (STEL), no NIOSH recommended exposure limits (REL), and no ACGIH TLV[1] for polyisocyanates although polyisocyanates in an aerosol form may be inhaled during spray-painting operations.[8-9] Some early reports suggest that polyisocyanates that are inhaled as an aerosol form are toxic and can cause occupational asthma during spray-painting operations.[2-4,10-11]

Columns 3 - 6 in Table I represent the results of analytical determinations performed over a period of ten days. All results obtained from a given analytical determination performed on a particular day for two side-by-side MAP/DMSO impingers are represented within a box. The average deviation from each other for the 12 data pairs is 0.38 mg/m3 (n=12) which is less than 0.51 mg/m3 (n=14) which is the pooled average deviation of a sample when run on different days. The results suggest that for these spray-painting operations, the average spatial variation in the concentration of polyisocyanate is small so long as the distance is less than 5 inches. The spatial variation is smaller than the temporal variations that are observed when running the same sample over a ten day period using the polyisocyanate-MAP HPLC procedure. The results seem to support the protocol which uses two impingers side-by-side for comparing sampling devices.

The table also lists the results for polyisocyanate when using the benchmark NIOSH Method 5521 (see last column). When comparing the two MAP results with those obtained from the MOP analysis, the results appear to be comparable except for the first and last sampling events.

References

1. American Council of Governmental Industrial Hygienists (ACGIH): *Threshhold Limit Values and Biological Exposure Indices for 1986-1987*. Cincinatti, Ohio: ACGIH, 1987.

2. Belin L., U. Hjortsberg, U. Wass: Life-Threatening Pulmonary Reaction to Car Paint Containing a Prepolymerized Isocyanate. Scand. J. Work Environ. Health 7: 310-312 (1981).

3. Malo J. L., G. Ouimet, A. Cartier, D. Levitz, R. Zeiss: Combined Alveolitis and Asthma Due to Hexamethylene Diisocyanate (HDI) with Demonstration of Crossed Respiratory and Immunologic Reactivities to Diphenylmethane Diisocyanate (MDI). J. Allergy Clin. Immunol. 72: 413-419 (1983).

4. Nielsen J., C. Sango, G. Winroth, T. Halberg, S. Skerfving: Systemic Reactions Associated with Polyisocyanate Exposure. Scand. J. Work Environ. Health 11: 51-54 (1985).

5. Occupational Safety and Health Administration: Diisocyanates (method 42). In OSHA Methods Manual. 1983. 42-1 - 42-39.

6. R. Streicher , J. Arnold, M. Ernst, and C. Cooper: Development of a Novel Derivatization Reagent for the Sampling and Analysis of Total Isocyanate Group in Air and Comparison of Its Performance with that of Several Established Reagents (Draft). Cincinatti, Ohio: NIOSH, 1995.

7. Janko, M., K. McCarthy, M. Fajer, J. van Raalte: Occupational Exposure to 1,6-hexamethylene Diisocyanate-Based Polyisocyanates in the State of Oregon, 1980-1990. Am. Ind. Hyg. Assoc. J.

53(5): 331-338 (1992).

8. Hardy, H., J. Devine: Use of Organic Isocyanates in Industry. Some Industrial Hygiene Aspects. Ann. Occup. Hyg. 22: 421-427 (1979).

9. Silk, S., H. Hardy: Control Limits for Diisocyanates. Ann. Occup. Hyg. 27: 333-339 (1983).

10. Vandenplas, O., A. Cartier, J. LeSage, Y. Cloutier, G. Perrault, L.C. Grammer, M.A. Shaughnessy, and J-L Malo: Prepolymers of Hexamethylene Diisocyanate as a Cause of Occupational Asthma. J. Allergy Clin. Immunol. 91(4): 850-861 (1993).

11. Tornling, G., R. Alexandersson, G. Hedenstierna, and N. Plato: Decreased Lung Function and Exposure to Diisocyanate (HDI) and (HDI-BT) in Car Repair Painters; Observations on Re-examination 6 Years after Initial Study. Am. J. Ind. Med. 17: 299-300 (1990).

GSRP FINAL REPORT TABLE OF CONTENTS i-xii

<u>APPENDICIES:</u>

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume7 is bound in two parts, 7A and 7B. Volume 7A contains #1-22. Volume 7B contains reports #23-45. The Table of Contents for Volume 2 is included in both parts.

This document is one of a set of 16 volumes describing the 1995 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3	Phillips Laboratory
4	Rome Laboratory
5A & 5B	Wright Laboratory
6	Arnold Engineering Development Center, Frank J. Seiler Research Laboratory, and Wilford Hall Medical Center
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10	Wright Laboratory
11	Arnold Engineering Development Center, Frank J. Seiler Research Laboratory, and Wilford Hall Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

DEVELOPMENT OF AN IN-VITRO CIRCULATORY MODEL WITH KNOWN RESISTANCE, INDUCTANCE, AND CAPACITANCE

Corey Offerdahl
Graduate Research Assistant
Cardiovascular Engineering Laboratory
Department of Electrical Engineering

Daniel L. Ewert, Ph.D.
Assistant Professor
Department of Electrical Engineering

North Dakota State University
Electrical Engineering 101
Fargo, ND 58105-5790

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1995

DEVELOPMENT OF AN IN-VITRO CIRCULATORY MODEL WITH KNOWN RESISTANCE, INDUCTANCE, AND CAPACITANCE

Corey D. Offerdahl
Graduate Research Assistant
Cardiovascular Engineering Laboratory
Department of Electrical Engineering
North Dakota State University of Agriculture and Applied Science

Daniel L. Ewert, Ph.D.
Assistant Professor
Department of Electrical Engineering
North Dakota State University of Agriculture and Applied Science

Abstract

An in-vitro (hydrodynamic) model of the circulatory system was developed. The model consisted of a pump, compliant tubing, and valves for resistance. The model is used to simulate aortic pressure and flow. These parameters were measured using a Konigsburg Pressure transducer and a Triton ART^2 flow probe. In addition, venous pressure and flow were measured on the downstream side of the resistance. The system has a known compliance, resistance, and inductance. Steady and pulsatile flow tests were conducted to determine the resistance of the model. A static compliance test was used to determine the compliance of the system. Inductance of the system was calculated using the equation $\rho L/A$. The aortic pressure and flow obtained from the hydrodynamic model can then be used to test the accuracy of parameter estimation models such as the 2-element and 4-element Windkessel models and the 3-element Westkessel model. Verifying analytical models used in determining total peripheral resistance (TPR) and systemic arterial compliance (SAC) is important because it provides insight into hemodynamic parameters that indicate baroreceptor responsiveness to situations such as changes in gravitational acceleration.

DEVELOPMENT OF AN IN-VITRO CIRCULATORY MODEL WITH KNOWN RESISTANCE, COMPLIANCE, AND INDUCTANCE

Corey D. Offerdahl Daniel L. Ewert

Introduction and Discussion of Problem

Pilots in military aviation experience transient changes in gravitational force that affect cardiovascular function. Severe changes in gravitational head-to-foot acceleration can lead to gravitational loss of consciousness (G-LOC). This occurs as blood pools in the lower extremities causing decreases in stroke volume and blood pressure. The lack of blood volume in the upper body limits the effectiveness of the baroreceptor's regulatory responses. Several circulatory analogs, such as the various Windkessel models [1, 2], have been developed to estimate various hemodynamic parameters that indicate the responsiveness and effectiveness of short term regulatory mechanisms such as total peripheral resistance (TPR) and systemic arterial compliance (SAC). Such analogs are only estimates, therefore it is important to know how accurate these estimates are to true hemodynamic values. The development of a hydrodynamic circulatory system with known resistance (TPR), compliance (SAC), and inductance (L, representing the inertial component of the fluid) lends insight to the accuracy of the Windkessel models.

Methodology

A hydrodynamic circulatory system was developed [3]. The system contained elements of known resistance, compliance, and inductance. The system consists of a pump (Cardiovascular Pulse Duplicator manufactured by Dynatek Laboratories, Inc.) which draws water from a reservoir through a mechanical atrial valve to a ventricle. The water was then pumped through a mechanical aortic valve, into a length of compliant tubing followed by a turnvalve resistor, and back into the reservoir. The total volume of fluid in the circulatory model was 442.5 ml. The input impedance of the setup was matched as closely as possible to that of the rhesus monkey (*mucaca mulatta*). Matching the artificial system's impedance to that of the rhesus monkey was accomplished by adjusting the resistance of the valve and by altering the length, thickness, and elastance of the compliant tubing. A rhesus model was used because of access to a supply of rhesus cardiovascular data. Simulated aortic pressure (AoP) and venous pressure (P_V) were measured using Konigsberg transducers (Konigsberg Instruments Inc, Pamona, CA). Simulated aortic flow (AoF) and venous flow(F_V) were measured using Triton ART^2 transit time flow meters (Triton Technology Inc, San Diego, CA). The pressure sensors and flow meters were contained in an in-line flow meter [4] that was developed at the laboratory and has been accepted for U.S. patent.

Pressure-Voltage calibration of the Konigsberg pressure transducers were done before each set of data were taken. Before closing the flow loop and performing experiments, a static compliance test was performed on the compliant tubing. The test consisted of injecting a known volume into the compliant tube and measuring the pressure change

(C = dV/dP) with the Konigsberg transducer. Data was taken in two identical sets consisting of nine runs. The order of the runs in each set was randomized. Heart rate (HR) and stroke volume were varied. The runs were completed as shown in Table 1.

Run #	HR (beats/min)	SV (ml)		Order of set 1 Run #	Flow Type		Order of set 2 Run #	Flow Type	
1	90	5		2	S	P	2	P	S
2	90	7		1	P	S	6	S	P
3	90	9		7	P	S	7	S	P
4	120	5		9	P	S	4	S	P
5	120	7		8	P	S	3	P	S
6	120	9		3	S	P	1	P	S
7	150	5		6	P	S	5	P	S
8	150	7		5	S	P	8	S	P
9	150	9		4	S	P	9	S	P

Table 1. Content and Order of Data Collection. Two sets of data were taken consisting of nine runs per set. Each run consisted of collection of steady flow (S) and pulsatile flow (P). The order of the runs were randomized as was the order of the flow type.

Each run was done using steady flow for resistance calculations (resistance was calculated as R= mean flow / mean pressure) and pulsatile flow for parameter estimation analysis. Steady flow was accomplished by installing a steady flow pump in the flow loop. Inductance was calculated using the equation $\rho = L/A$ (ρ = density of fluid, L = length of tube, A = cross-sectional area). It was assumed that compliance and inductance stayed constant from initial values. Flow calibration on the Triton flow probes was done after every fourth run. Data were recorded using an A/D card (Data Translations 7805), that sampled at 125 Hz and used anti-aliasing filters with a cutoff frequency of 60 Hz, and stored on a PC (Zeos 486 Pentium 90). Analysis of the signals was done using MATLAB (The Mathworks Inc., Natick, MA).

Results

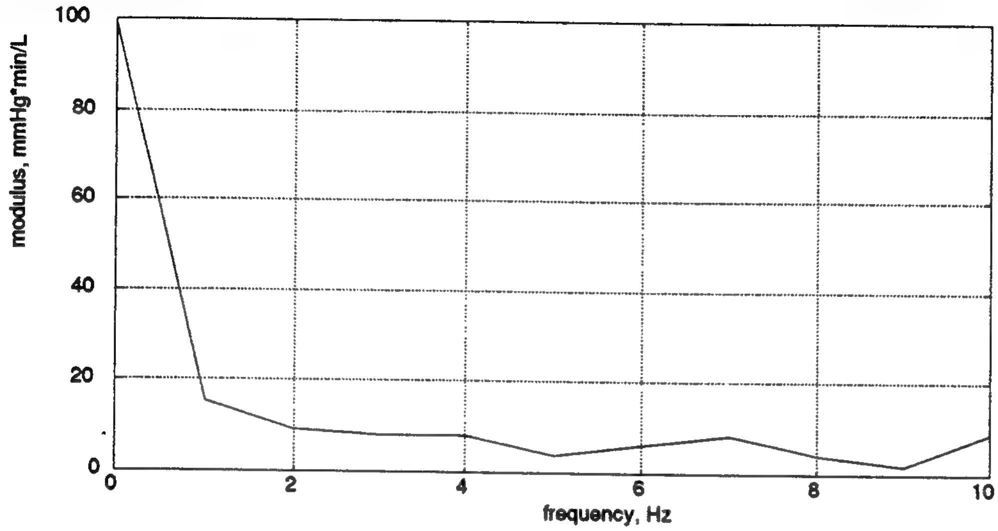
Impedance Matching to Rhesus Monkey

The impedance of the hydrodynamic circulatory system was matched to that of the rhesus monkey. In order to match the impedance, rhesus aortic pressure and flow data were analyzed. The Fast Fourier Transform of AoP and AoF were taken for one beat. Input Impedance was found by taking:

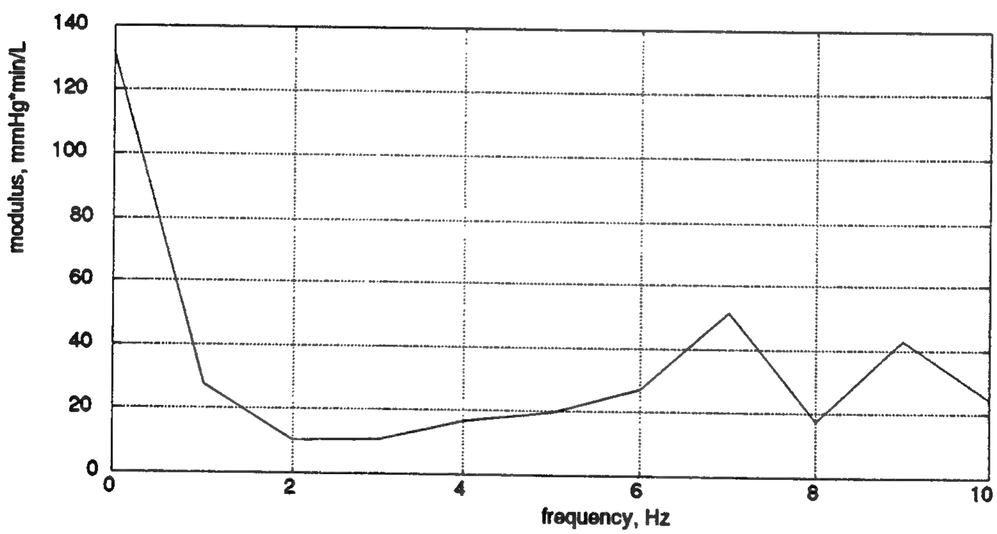
$$\mathrm{FFT(AoP)/FFT(AoF)} \qquad (1)$$

This was repeated for several beats to find a range of input impedance magnitude and phase.

By adjusting the resistance of the hydrodynamic system and by using different compliant tubes, a best fit model was determined. It was found that input impedance magnitude could not be matched as closely as phase. The shape of the magnitude plots are similar, however the magnitude of the in-vitro impedance is approximately 1.3 times that of physiological data (Figure 1). The phase of the hydrodynamic system matched closely to physiological phase at both the fundamental frequency and at the crossover point (Figure 2).

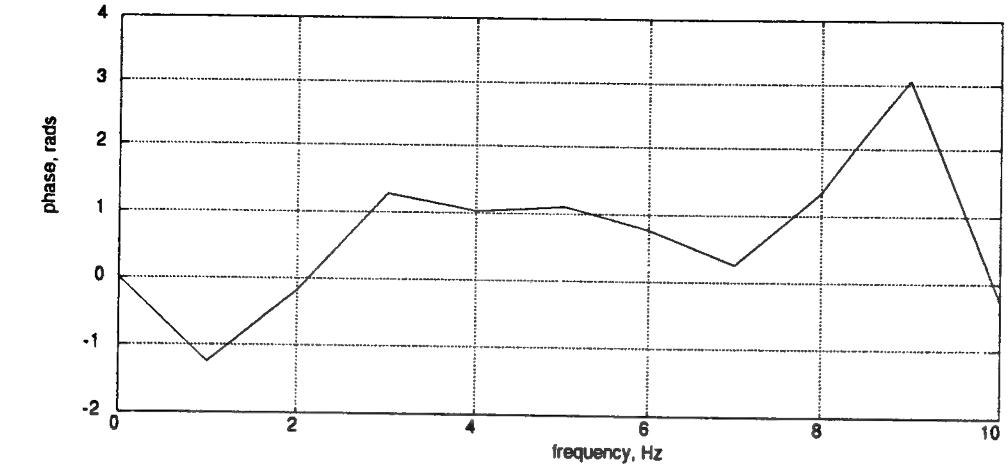


(a)

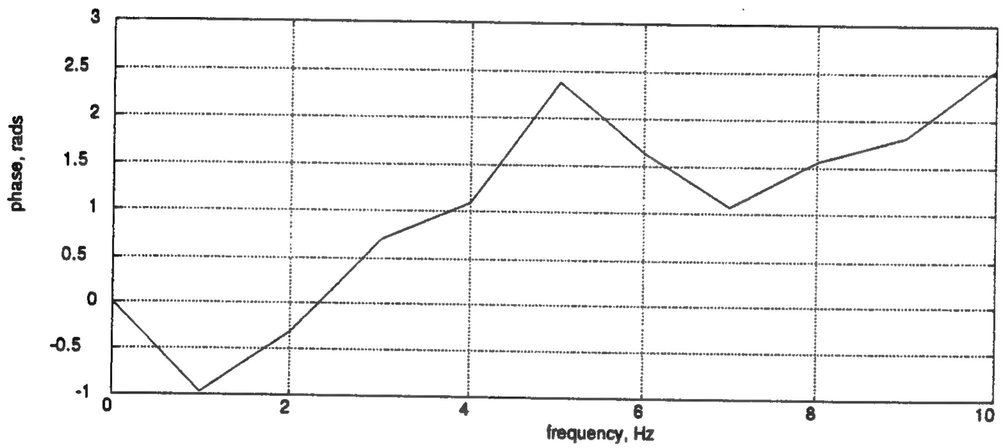


(b)

Figure 1. (a) Modulus of the input impedance of the Rhesus Monkey. (b) Modulus of the input impedance of the hydrodynamic circulatory model.



(a)



(b)

Figure 2. (a) Phase of the Aortic input impedance of a Rhesus monkey. (b) Phase of the input impedance of the hydrodynamic circulatory model.

Static Compliance Test

Static compliance was measured by injecting a known volume into the compliant tubing and measuring the pressure change that occurs.

$$C = dV/dP \qquad (2)$$

Static compliance tests were done before and after each run. The results are shown in Figure 3. It was assumed that the compliance of the tube remained constant during pulsatile flow. The average compliance was found to be 0.093 mmHg in the operating region (~ 50-150 mmHg) of the hydrodynamic model.

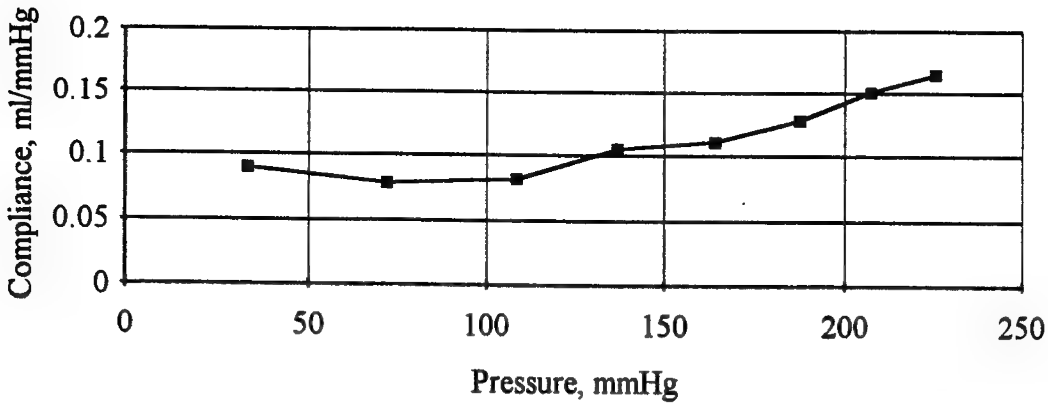


Figure 3. Result of the static compliance test. Compliance is given as a function of pressure. Compliance remains fairly constant in the operating range of the hydrodynamic model (50-150 mmHg). The average compliance in this region is approximately 0.093 ml/mmHg.

Steady Flow Resistance Test

Resistance was calculated using steady flow by the following equation:

$$R = \text{Mean Pressure/Mean Flow} \qquad (3)$$

Flow was adjusted from zero to two liters per minute. The results of the steady flow resistance test are given in Figure 4. It was assumed that resistance of the system did not change under pulsatile flow.

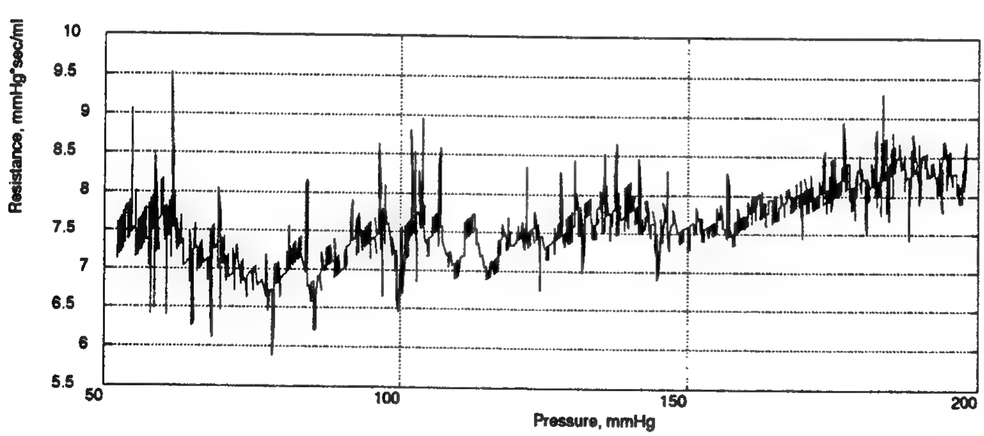


Figure 4. Resistance of the hydrodynamic circulatory model is shown as a function of pressure. The average resistance is approximately 7.836 mmHg•sec/ml.

Inductance

Inductance was calculated using the equation:

$\rho L/A$ (4) ρ = density of the fluid (1 g/cm^3 for water)

L = length of the tube through which flow occurs

A = cross-sectional area of the tube

Inductance was assumed to remain constant during pulsatile and steady flow. The total inductance of the hydrodynamic model is 78.25 g/cm^4.

Representative Data From the In-Vitro Model and The Rhesus Monkey

Figure 5 shows examples of the data obtained from the hydrodynamic model and from the rhesus monkey.

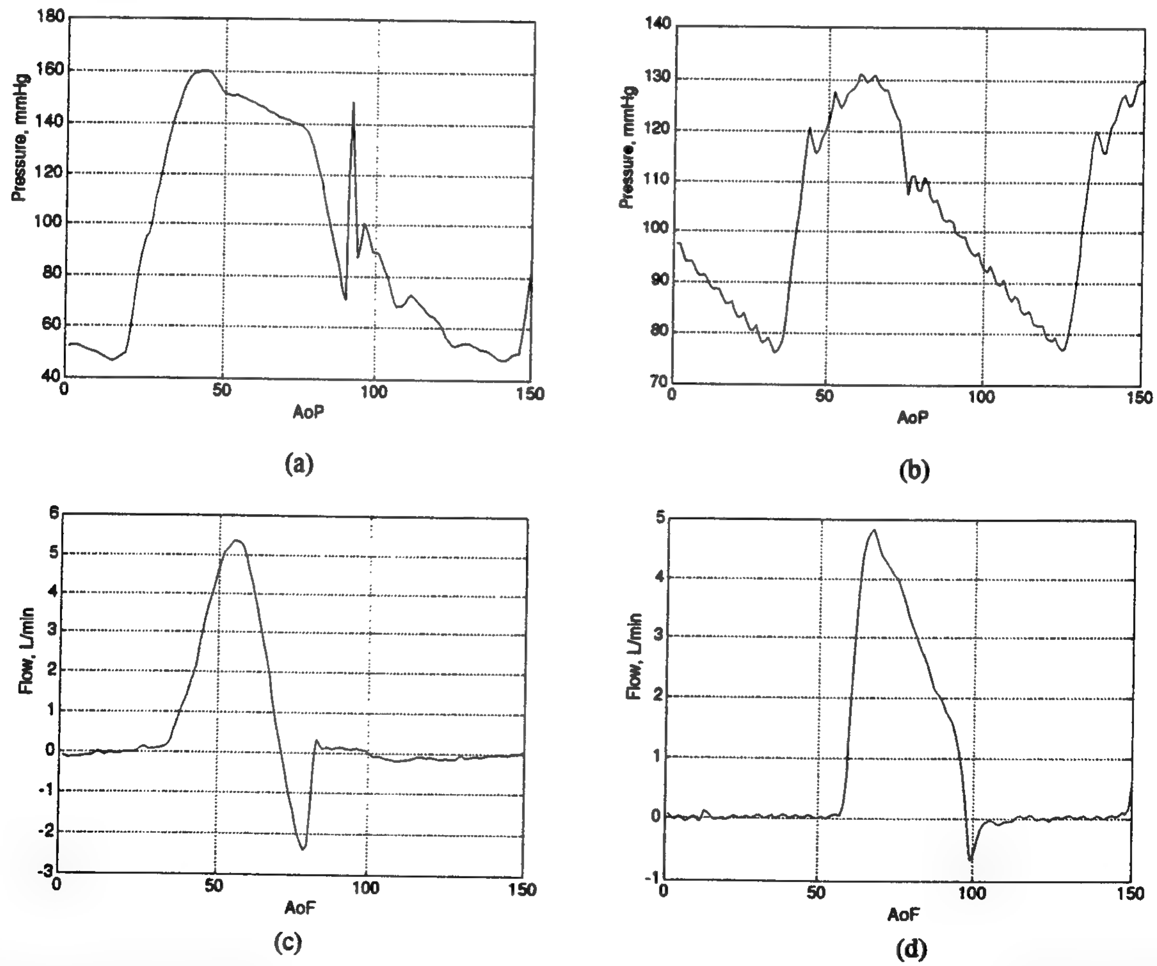


Figure 5. (a) Aortic Pressure from the hydrodynamic model. (b) Aortic Pressure from a Rhesus monkey. (c) Aortic Flow from the hydrodynamic model. (d) Aortic Flow from a Rhesus monkey.

Figure 5 shows that the pulse pressure of the hydrodynamic model (Figure 5(a)) is greater than that of the Rhesus monkey (Figure 5(b)). The pressure spike near the dicrotic notch is an artifact that corresponds to the closure of the aortic valve (Figure 5(a)). It also shows that the negative aortic flow component (Figure 5(c)) is much larger than that of the Rhesus monkey (Figure 5(d)). This occurs due to backward flow through the mechanical aortic valve, something that does not occur with normal physiological valves.

Discussion and Conclusion

An in-vitro (hydrodynamic) circulatory model was developed. The input impedance of the model was matched to the input impedance of the rhesus monkey. It is shown in Figure 1 and Figure 2 that the modulus and phase of the impedance match well with that of the rhesus monkey. The modulus of the hydrodynamic model is higher than that of the rhesus monkey. This may be due to the high aortic pulse pressure (Figure 5) that occurs in the model. Phase (Figure 2) compares favorably between the animal and hydrodynamic model. The phase crossover frequency is nearly the same in both cases. The animal circulatory system is a bit more capacitive than the model as the phase is lower at the fundamental frequency (1 Hz). Matching the impedances will allow better comparison between animal and hydrodynamic data when it is applied to the Windkessel models.

The model does have some limitations due to the assumptions made. The assumption that static compliance remains constant under pulsatile flow is limited by hysteresis that occurs in the non-ideal compliant tubing. The assumption that the inertia or inductance remains constant is limited by the change in length and cross-sectional area of the tubing due to the pulsatile nature of the flow. In addition, the amount of backflow through the mechanical aortic valve may cause error in parameter estimation. This could be alleviated by use of a better quality artificial valve.

Hopefully, the hydrodynamic model and the data derived from it will help determine how good the Windkessel models are at estimating vascular parameters. In the event that the estimates do not match the known resistance, compliance, and inductance, new insight may be found that allows adaptation of the Windkessel models to estimate these parameters more accurately. Such modifications may include the addition of wave reflections to them.

References

[1] Self, David A., Daniel L. Ewert, Richard D. Swope, et. al, "Beat-to Beat Determination of Peripheral Resistance and Arterial Compliance During +Gz Centrifugation," *Aviat. Space Environ. Med.* vol. 65, no.5, pp. 396-403.

[2] Essler, Steven J., "Estimation of Arterial Vascular Parameters for Transient and Steady Beats," M.S. Thesis, North Dakota State University, 1995.

[3] Schaub, J. D., S.C. Koenig, C.D. Offerdahl, D.L. Ewert, R.D. Swope, "Cardiovascular Hemodynamics Model," Med. & Biol. Eng. & Comp. (Submit Oct. 95).

[4] Koenig, S.C., J.D. Schaub, D.L. Ewert, R.D. Swope, "In Line Pressure-Flow Module For CV Hydrodynamics Modeling and Biosensor Evaluation." Med. & Biol. Eng. & Comp. (Submit Oct. 95).

A STUDY OF THE RECOVERY OF TRICHLOROBENZENE FROM A CATIONIC ENHANCED SORPTION ZONE IN COLUMBUS AQUIFER MATERIAL USING A NONIONIC SURFACTANT

Angela J. Parlier
Graduate Student
Department of Civil Engineering

Auburn University
238 Harbert Engineering Center
Auburn, AL 36849-5337

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

October 1995

A STUDY OF THE RECOVERY OF TRICHLOROBENZENE FROM A CATIONIC ENHANCED SORPTION ZONE IN COLUMBUS AQUIFER MATERIAL USING A NONIONIC SURFACTANT

Angela J. Parlier
Graduate Student
Department of Civil Engineering
Auburn University

Abstract

The viability of the recovery of trichlorobenzene (TCB) from a cationic surfactant enhanced sorption zone in Columbus aquifer material using a nonionic surfactant was studied. A simulation of the cationic enhanced sorption zone was created by treating Columbus aquifer material. TCB was introduced into the column and shown to be retarded in the enhanced sorption zone. When flushed with a nonionic surfactant at a concentration greater than the CMC, the TCB will partition into the micelle phase and can be recovered.

Introduction

Surface active agents, or surfactants, have been studied recently as a means of remediating groundwater contaminated with hydrophobic organic contaminants (HOCs). Some surfactants have been shown to increase the aqueous solubility of HOC's resulting in an increase in the recovery of the contaminant (1-4). When coupled with a pump-and-treat remediation scheme, solubility enhancement may be an effective method of contaminant recovery.

Surfactants are amphiphillic molecules with a hydrophilic polar head and a hydrophobic nonpolar organic carbon tail. This characteristic makes it possible for surfactant monomers to adsorb at different phase boundaries. Surfactants also have the ability to form micelles when the aqueous concentration reaches the critical micelle concentration (CMC). At the CMC, hydrophobic tails of the surfactant monomers begin to interact, forming mobile molecular structures which can act as a partitioning phase for HOC's (2,5-7). Figure 1 shows the general interaction of surfactant monomers between the aqueous and solid phase at different aqueous surfactant concentrations.

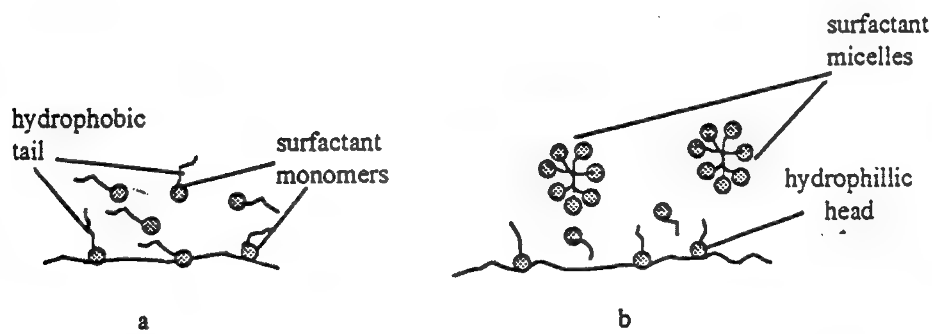


Figure 1. Surfactant behavior as monomers and micelles a) below the CMC, and b) above the CMC

Cationic surfactants can artificially increase the organic carbon (OC) content of some soils, thereby creating an enhanced sorption zone when injected into an aquifer (8-11). At aqueous concentrations less than the CMC, the cationic surfactant monomers will bind to the soil at the sorption sites by cation exchange. HOC's can then sorb to the hydrophobic carbon tail of the cationic surfactant. This decreases the mobility of the aqueous phase HOC which prevents further transport of a contaminant.

The research presented in this report deals with the recovery of TCB from a cationic surfactant-enhanced sorption zone using the nonionic surfactant IGPAL CO-730 (Rhone-Poulenc, Inc.). Previous studies have shown that a nonionic surfactant increases solubility of HOC's when the aqueous concentration of the surfactant is above the CMC by providing a mobile partitioning phase for the contaminant. The objective of this study is to determine the viability of a remediation scheme where HOC's are first immobilized within a cationic surfactant enhanced-sorption zone, then re-mobilized using a nonionic surfactant flush.

Materials and Methods

Batch Experiments

Experimental batch sorption studies were conducted on a natural aquifer material acquired from Columbus Air Force Base, Mississippi. The cationic surfactant, hexadecyltrimethylammonium chloride (HDTMA), was used to enhance the sorptive capacity of the soil. The nonionic surfactant, CO-730, was used to remobilize sorbed HOC's. The different systems to be prepared and studied were: CO-730, soil, and water; CO-730, HDTMA, soil, and water; and CO-730, HDTMA, TCB, soil, and water. These systems were to be studied in order to determine the partitioning behavior of the nonionic surfactant in natural aquifer material and in the presence of the cationic surfactant-treated soil; and the partitioning behavior of the HOC in a cationic surfactant-enhanced sorption zone.

In an actual field application of this proposed remediation strategy, a cationic surfactant enhanced sorption zone would be created by injecting a cationic surfactant solution into the saturated zone of an aquifer (8-9). For the experiments performed in this study, the aquifer material was simply treated as follows: 2 kilograms of the Columbus aquifer material (< 2 mm fraction) was placed in a large tray. An aqueous solution of 6720 mg/L HDTMA was placed in the tray and the contents were stirred in order to homogenize the material. The mixture was allowed to stand for 48 hours in order to ensure that equilibrium between the soil and surfactant was reached. The soil was then rinsed well with deionized water and allowed to settle. After settling, the liquid was poured off, centrifuged, and the fine particles were recovered. The aquifer material was then oven dried and rehomogenized. The organic carbon content was measured before

and after treatment using a LECO organic carbon analyzer The organic carbon content of the untreated soil was approximately 0.07% OC while the organic carbon content of the treated soil was approximately 0.33% OC. This is a good representative value for a fully developed enhanced sorption zone in Columbus aquifer material (9).

The CMC of the nonionic surfactant in the presence of treated and untreated soil was determined from surface tension measurements using the drop-weight method described by Adamson (11).

Batch sorption equilibrium isotherms were used to determine the sorption characteristics of CO-730 on soil both treated with HDTMA and untreated. Batch vials were prepared by mixing ~5 grams of aquifer material with ~20 mg/L of CO-730 solution. These samples were placed on a roller drum and gently equilibrated for 24 hours at room temperature (21^{0} C). They were then centrifuged for 15 minutes at 1500 RPM. The supernatant was then analyzed for CO-730 concentration using UV-Visible Spectrophotometer (Varian, Inc.) at 274 nm wavelength. The partitioning was determined from the difference of the measured aqueous CO-730 mass the known initial mass to give the sorbed mass.

TCB sorption onto treated and untreated soil was measured. A solution containing a known amount of aqueous TCB was injected into batch vials containing the aquifer material. The vials were treated in the same manner as stated before. After equilibrium, the aqueous TCB was measured using gas chromatography.

Column Experiments

One-dimensional transport studies were performed in a column in order to determine the breakthrough curves of the different systems. The stainless steel column was 25 cm in length and 2.1 cm inner diameter. Two 2 μm stainless steel frits were used to contain the soil within the column. The column was weighed and then aquifer material was placed in the column by slowly pouring in the soil while gently vibrating the column in order to ensure uniform packing. The newly packed column was flushed with nitrogen gas to purge all air from the system. The column was then flushed with a solution of 0.02% sodium azide, NaN3 for ~2 hours. This solution controlled microbial growth as well as providing background ionic strength. The three column experiments that were done were CO-730 in untreated soil, CO-730 in treated soil, and TCB followed by CO-730 in treated soil.

A generalized sketch of the column system is shown in Figure 2. All aqueous solutions were kept in Pyrex bottles which were connected to an ISCO Model 2360 Gradient Programmer and Pump which regulated the flow rate at 0.25 ml/min. A syringe pump delivered the TCB solution at a constant flow rate of 0.25 ml/min. All aqueous solutions contained 0.02% NaN_3 in order to prevent unwanted microbial activity and for ionic strength. The effluent from the column was collected by a Gilson 212B Liquid Handler. All samples were collected over a 6 minute interval.

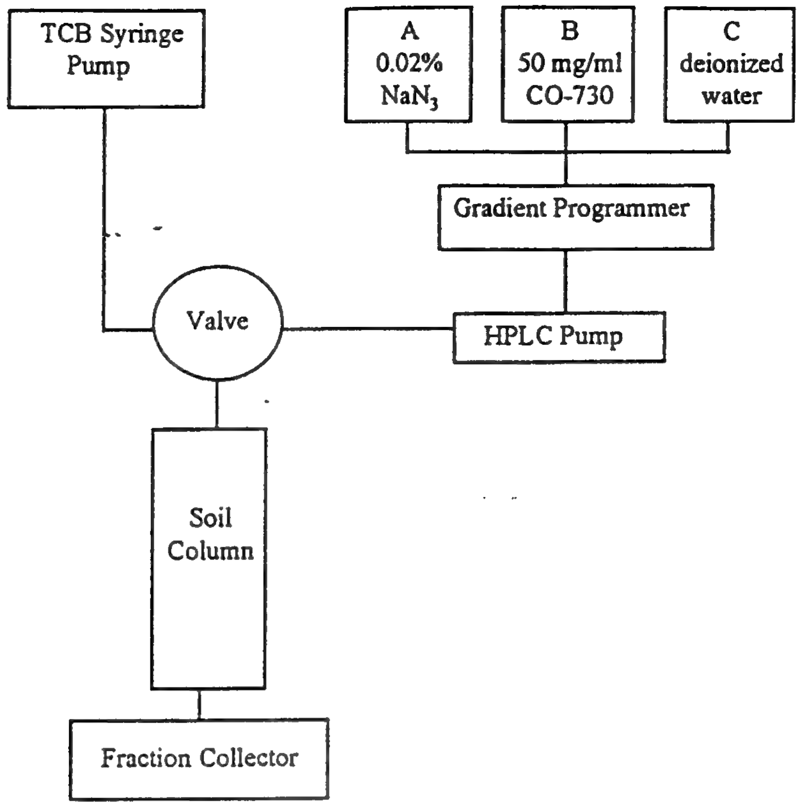


Figure 2. Column System Setup

Three column experiments were performed and a new column was used for each experiment. Longitudinal dispersivity was estimated from the breakthrough of a conservative tracer, NaCl (Figure 3). The first column experiment utilized untreated Columbus aquifer material and the breakthrough curve of CO-730 was measured. The second column utilized treated aquifer material and again, the breakthrough curve of CO-730 was measured. The final column experiment utilized treated aquifer material. A 19 mg/L TCB solution was injected into the column for 5 hours followed by a flush of a 0.02% NaN_3 solution for 88 hours. The column was then flushed with a 50 mg/l CO-730

solution for approximately 15 hours and the breakthrough of both the TCB and the nonionic surfactant was measured.

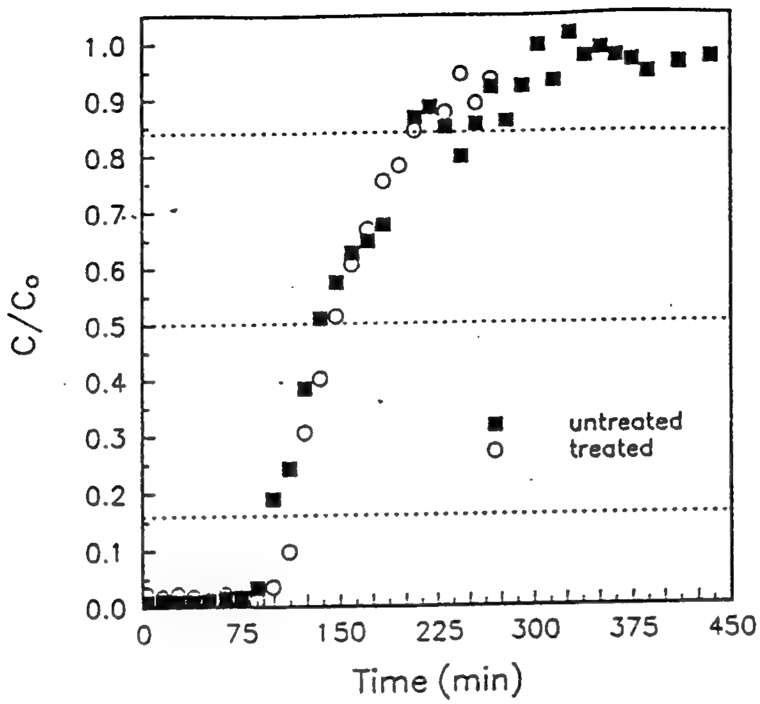


Figure 3. Breakthrough curve of the conservative tracer, NaCl.

Results and Discussion

Batch Experiment Results

A 2-term, nonlinear Langmuir isotherm shown in EQUATION 1 (12) was used to describe the sorption of CO-730 in treated soil and soil treated with HDTMA. The constants were fitted using a numerical code.

$$Cs = \frac{b1K1Ca}{1+K1Ca} + \frac{b2K2Ca}{1+K2Ca} \qquad \text{(Equation 1)}$$

This equation is a reasonable fit to the experimental data where b1 and b2 are 9.705355 and 8.03096 mg CO-730 sorbed/gms of soil respectively, K1 and K2 are 9.510213 and 0.1451367 ml solution/mg CO-730 in the aqueous phase respectively, Ca is the adsorbed

concentration, and Ca is the aqueous concentration of CO-730. Figure 4 shows the sorption isotherms.

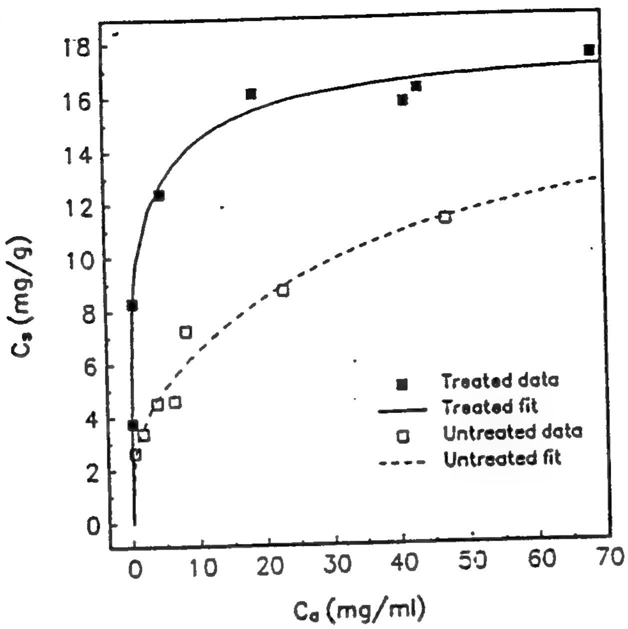


Figure 4. Two-term Langmuir Sorption Isotherm

Surface tension measurements provided the CMC of the different surfactant systems. It has been found that the HDTMA desorption rate is constant (13). Figure 5 shows that the CMC of CO-730 and water is much higher by an order of magnitude than when HDTMA is present in the aqueous phase. It is also important to notice that the presence of soil does not appear to have an effect on the formation of micelles. HDTMA in the aqueous phase appears to be the governing factor in the formation of micelles in the different systems.

The apparent solubility of TCB in the presence of CO-730 and in the presence of CO-730 and HDTMA (Figure 6) are shown to be different by an order of magnitude. The apparent solubility of TCB increases with an increase in CO-730. Segment 1 gives

the Kmn value (partitioning rate between the TCB and the monomers) in the distribution equation (Equation 2):

$$Kd = Cs/ Ca \qquad \text{(Equation 2)}$$

where Cs is the sorbed concentration of the TCB, Ca is the aqueous concentration of TCB, and Kd is the distribution coefficient (14). Segment 2 give the Kmc value (partitioning rate between the TCB and the micelles) in the distribution equation. The break in the line is the CMC and shows where the partitioning of TCB into the micelles begins. These values describe the partitioning relationship between the TCB and the surfactant systems.

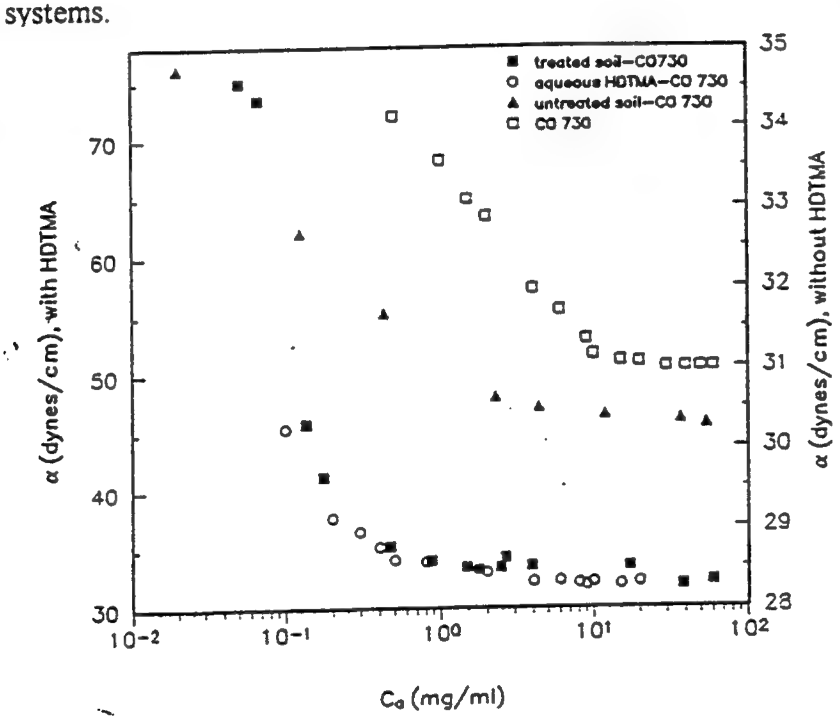


Figure 5. CMC determination from surface tension experiments.

Figure 7 shows the different Kd values as a function of CO-730 concentration and indicates how the linear equilibrium partitioning coefficient depends on aqueous nonionic surfactant concentration. Figure 8 is a 10th order polynomial fitted line that was done with a modeling program showing the variation of the Kd value as a function of aqueous

surfactant concentration. Figure 8 is a 10th order polynomial fitted line that was done with a modeling program showing the variation of the Kd value as a function of aqueous CO-730. At the peak of the curve is the CMC of the nonionic surfactant and shows the highest Kd value.

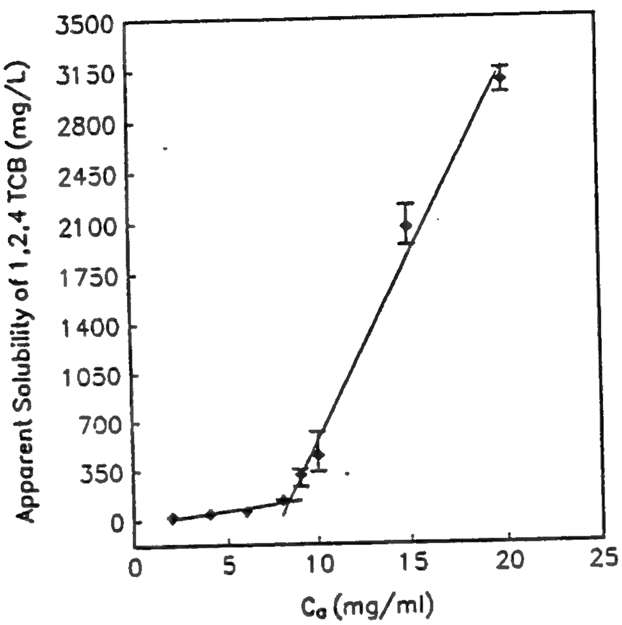


Figure 6. Apparent solubility of TCB in the presence of the different surfactant systems.

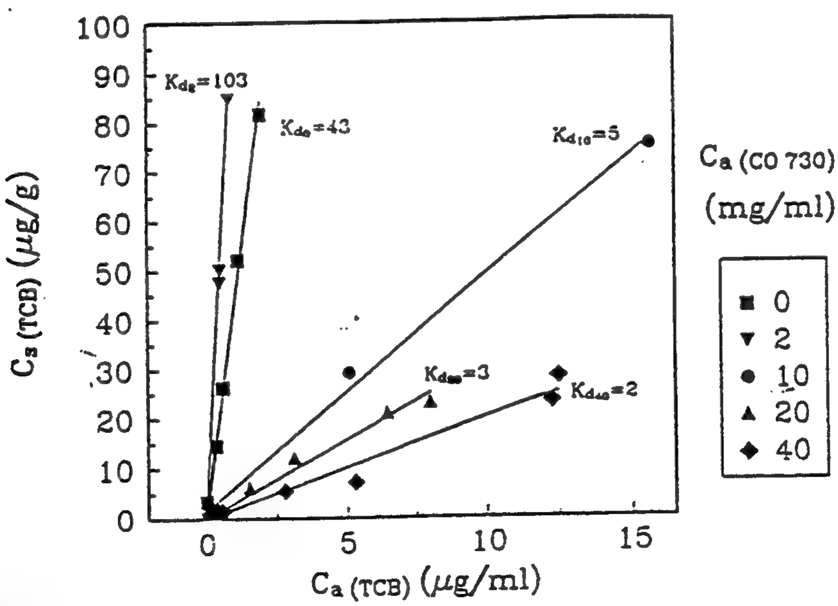


Figure 7. Kd values as a function of nonionic surfactant concentration.

Column Experiment Results

Figure 9 is the breakthrough curve of the nonionic surfactant through untreated aquifer material. By assuming an equilibrium model, there is a lot of variation from the experimental data. The kinetic data, however shows a much better fit to the experimental suggesting that the nonionic surfactant transport is nonequilibrium process where the rates of sorption and desorption are different and not necessarily instantaneous (15). Similarly, the movement of the nonionic surfactant through treated aquifer material (Figure 10) shows that there are kinetic effects.

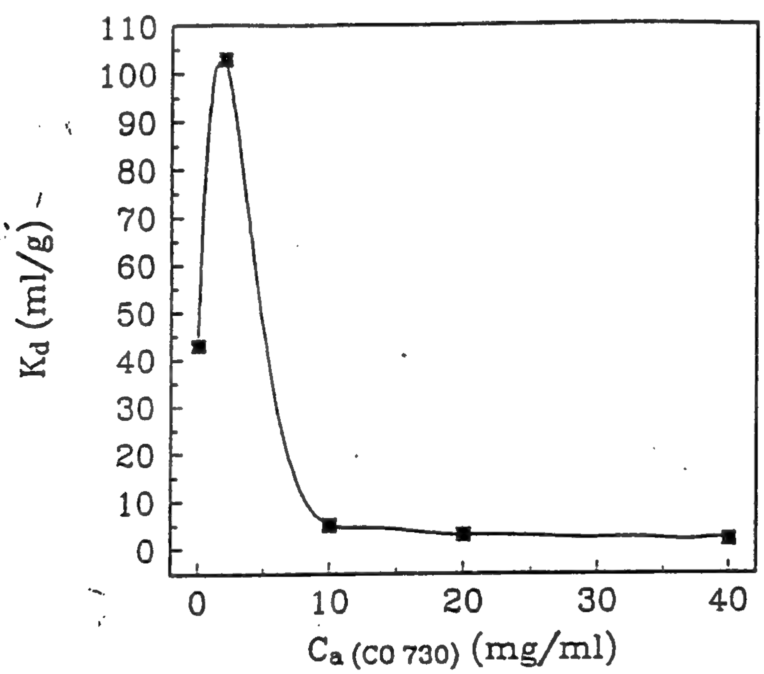


Figure 9. Breakthrough curve of the nonionic surfactant through untreated aquifer material

because less nonionic surfactant will be needed for remediation. The second conclusion is that kinetic effects in transport dictate distribution of nonionic surfactant. However, TCB transport is not kinetic but is dependent on aqueous phase nonionic surfactant. I hope to show that when a soil column containing treated soil is injected with TCB, the HOC's become basically immobile. When the nonionic surfactant is introduced into the column, the TCB will be recovered based on the strong partitioning of TCB into the micellephase as opposed to a solid phase.

Literature Cited

(1) Abdul, Abdul S.; Gibson, Thomas L.. *Laboratory Studies of Surfactant-Enhanced Washing of Polycholorinated Biphenyl from Sandy Material.* Environ. Sci. Technol. **1991**, 25, 4.

(2) Edwards, David A.; Liu, Zhongbao; Luthy, Richard G. *Surfactant Solubilization of Organic Compounds in Soil/Aqueous Systems.* J. of Environ. Eng. **1994**, 120,1.

(3) Edwards, David A.; Luthy, Richard G.; Liu, Zhongbao. *Solubilization of Polycyclic Aromatic Hydrocarbons in Micellar Nonionic Surfactant Solutions.* Environ. Sci. Technol. **1991**, 25, 1.

(4) Pennell, Kurt D.; Jin, Minquan; Abriola, Linda M.; Pope, Gary A.. *Surfactant enhanced remediation of soil columns contaminated by residual tetrachloroethylene.* J. or Contaminant Hydrology. **1994**, 16.

(5) Sun, Shaobai; Inskeep, William P.. *Sorption of Nonionic Organic Compounds In Soil-Water Systems Containing a Micelle-Forming Surfactant.* Environ. Sci. Technol. **1995**, 29, 4.

(6) Deitsch, James J.; Smith, James A. *Effect of Triton X-100 on the Rate of Trichloroethene Desorption from Soil to Water.* Environ. Sci. Technol. **1995**, 29, 4.

(7)) Edwards, David A.; Liu, Zhongbao; Luthy, Richard G. *Experimental Data and Modeling for Surfactant Micelles, HOC's, and Soil.* J. of Environ. Eng. **1994**, 120, 1.

(8) Brown, Michael John. *Enhanced Retardation of Organic Contaminants on Aquifer Material Modified with Cationic Surfactants.* Diss. U. of Waterloo. 1993.

(9) Burris, David R.; Antworth, Christopher P. *In situ modification of an aquifer material by a cationic surfactant to enhance retardation of organic contaminants.* J. of Contaminant Hydrology. **1992**, 10.

(10) Wagner, Julia; Chen, Hua; Brownawell, Bruce J.; Westall, John C. *Use of Cationic Surfactants To Modify Soil Surfaces To Promote Sorption and Retard Migration of Hydrophobic Organic Compounds.* Environ. Sci. Technol. **1994**, 28, 2.

(11) Adamson, Arthur. *Physical Chemistry of Surfances*, 4th ed. John Wiley & Sons. 1982. pp 20-23.

(12) Sposito, Garrison. *The Surface Chemistry of Soils.* New York: Oxford University Press, 1984.

(13) Westall, John C.; Hatfield, Julia; Chen, Hua; Brownawell, Bruce J. ***The Use of Cationic Surfactants to Modify Aquifer Materials to Reduce the Mobility of Hydrophobic Organic Compounds - A Study of Equilibrium and Kinetics.*** U. S. EPA. Robert S. Kerr Environ. Res. Laboratory, 1994.

(14) Freeze, R. Allen; Cherry, John A. *Groundwater.* New Jersey: Prentice-Hall, Inc., 1979.

(15) Hayworth, Joel S.; Burris, David R. *Modeling Cationic Surfactant Transport in Porous Media.* accepted April 5, 1995 for publication in Ground Water.

THE SORPTION OF HUMIC ACIDS ONTO AQUIFER MATERIAL SOLIDS

Michael J. Piana
Graduate Student
Department of Chemistry

California State University, Northridge
18111 Nordhoff Street
Northridge, California 91325

Final Report for:
Summer Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office Of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1995

THE SORPTION OF HUMIC ACIDS ONTO AQUIFER MATERIAL SOLIDS

by

Michael J. Piana

Graduate Student
Department of Chemistry
California State University, Northridge

Abstract

A study was made of the sorption of commercial Aldrich humic acid and IHSS Suwannee River humic acid onto three natural aquifer material solids from Barksdale AFB, Louisiana; Blytheville AFB, Arkansas and Columbus AFB, Mississippi . The interactions follow trends found in other humic acid-mineral sorption studies. Increase in pH caused sorption of humic acid on each sediment to decrease. Humic substances sorb more strongly to sediments with considerable surface areas, and surface iron coupled with substantial silt and clay percentages relative to sand. Humic acids with similar O/C ratios and aromatic content had similar sorption capacities on the aquifer material. Ionic strength effected sorption, the concentration and type of ionic strength buffer used governed the extent to which humics sorbed. Phosphate buffer competes with the humic acid for sorption onto the aquifer material solid. Perchlorate buffer tends to alter the structure of the humic acid causing it have different sorption capacities on the aquifer solids. Sorption was greater on the aquifer solids at I = 0.1 M than at I = 0.005 M. All isotherms were evaluated for Freundlich, Langmuir and linear fits. The sorption data favored Freundlich-type statistical treatment.

THE SORPTION OF HUMIC ACIDS ONTO AQUIFER MATERIAL SOLIDS

Michael J. Piana

Introduction

The dark colors observed in many soils, sediments, and natural waters are the characteristic of humic substances. Humic substances are compounds formed from the biological degradation of plant and animal residue and from the synthetic activities of microorganisms. Humic materials are some of the most powerful metal-binding agents among natural organic substances and can form water-soluble and insoluble complexes with metal ions and hydrous oxides (Tipping, 1981). The structure of humic substances are extremely complex, humic materials are high molecular weight (1000-500,000 daltons) compounds comprised of numerous polyfunctional groups. including; carboxyl, phenolic, alcoholic hydroxyl, quinone and ketonic carbonyl, amino, and sulfhydryl groups. As a result of their large size and chemical complexity, humic materials have been compared to biomolecules in structure and properties. However, in contrast to biomolecules, humics are quite complex and can differ structurally from one molecule to the next. Sposito (1984) summarized the principle characteristics of humic substances that influence their chemical reactivity: (1) polyfunctionality, the existence of a variety of functional groups and a broad range of functional reactivity, representative of a mixture of interacting heteropolymers (2) macromolecular charge, the development of polyelectrolyte character with the resultant effect on functional group reactivity and molecular conformation (3) hydrophilicity, the tendency to form strong hydrogen bonds with water molecules solvating polar functional groups and (4) structural lability, the capacity to associate intermolecularity and to change molecular confirmation in response to changes in pH, pE values, electrolyte concentration, and functional group binding.

Humic substances provide important sources of both dissolved and colloidal organic ligands and therefore, can influence the bioavailability of trace metals in soils, sediments and aquatic systems. (Sposito, 1984; Tipping and Cooke, 1981; Davis and Bhatnagar, 1995) The interaction between a metal ion species and a humic substance is best classified as a "sorption reaction" varying from purely electrostatic to strongly covalent. (Sposito 1984). The bond that forms between the humic and metal species and surface functional groups is often not always clearly defined. Due to the polyfunctional nature of humics, the metal ion may attach to the ionized peripheral groups of the humic material or react with the functional groups creating inner-sphere and outer-sphere complexes.

Humic substances can sorb to mineral surfaces containing hydroxylated Al, Fe, or Mn sites thereby changing the mineral surface properties such as rendering the hydrophilic mineral surfaces hydrophobic, and more capable of sorbing metal species (Tipping and Cooke, 1982; Murphy et. al., 1990; and Zachara, 1994). Surface bound humic substances increase the adsorption of certain metal cations on single mineral solids by contributing additional, potentially high affinity, complexation sites. For metals forming strong complexes with humic substances, humic substances enhance metal ion binding to Fe and Al oxides at lower pH (e.g.,<6) by cosorption, but decrease metal ion sorption at higher pH by formation of nonsorbing aqueous complexes (Tipping et al.,1981; Tipping and Cooke, 1981; Davis and

Bhatnagar, 1995; and Murphy et. al., 1990). Consequently, particle-associated humic substances may form coatings on minerals, given that natural dissolved and colloidal organic material bind to a variety of oxide (Al and Fe) and layer silicate surfaces that are common in aquifers. (Tipping, 1981; Tipping and Cooke, 1982; Stumm, 1987; Murphy et al.,1992).

The structure of dissolved humic substances is largely determined by the groundwater pH, ionic strength, and the presence of divalent cations (Tipping, 1981; Tipping and Cooke, 1982; Christman and Gjessing, 1983; Murphy et. al., 1990; Murphy et. al., 1992) At high ionic strength, low pH, or in the presence of cations, the charge repulsion between adjacent carboxyl or hydroxyl groups on the humic substance is neutralized, resulting in a coiled configuration (Tipping and Cooke, 1982; Tsutsuki and Kawasuka, 1983) . At the mineral/solution interface the coiled configuration can result in fewer attachment points between the humic carboxyl and mineral surface hydroxyl groups and an extension of a shear plane outward from the mineral surface. Conversely, at low ionic strength, high pH and in the absence of cations, the humic substance may adopt a more open configuration in solution, resulting in more attachment points upon sorption to the mineral surface (Tipping and Cooke, 1982; Tsutsuki and Kawasuka, 1983; Murphy et al., 1992; Zachara et.al., 1994).

The humic material is susceptible to configuration changes with different ionic strength buffer. Reagents such as perchlorates and phosphates for example, have an effect on the sorptive behavior of the humic material to the aquifer solids. In some cases, the buffering solution may be involved in direct competition with the humic materials for specific sites on the aquifer material. It is believed that phosphates (PO_4^-) compete directly with the humic material for the Al-OH and Fe-OH sites. The extent of composition on the physical properties of the aquifer material and the concentration of the buffer itself. (Tipping, 1981; Tipping and Cooke, 1982; Tsutsuki and Kawasuka, 1983; Murphy et. al., 1992).

Published data for humic substance sorption onto mineral surfaces suggest that sorption tends to follow a Langmuir-type isotherm curve fit (Tipping, 1981; Murphy et. al., 1990). The Langmuir isotherm is based on the concept that a solid surface poses a finite number of sorption sites. When all the sorption sites are filled, the surface will no longer sorb solute from solution thus a sorption plateau is reached when all sites are filled (Fetter, 1993). The form of the Langmuir sorption isotherm is:

$$C/C^* = 1/\alpha\beta + C/\beta$$

where:

C = initial amount substance present

C^* = amount of material sorbed at the end of reaction

α = an absorption constant related to the binding energy (l/mg)

β = the maximum amount of solute that can be absorbed by the solid (mg/kg)

If C/C^* is plotted versus C, the data will be linear with maximum ion sorption, β, as the reciprocal of the slope of the line and the binding energy constant, α, as the slope of the line divided by the intercept.

A more general or true equilibrium isotherm is the Freundlich sorption isotherm. This is defined by the nonlinear relationship:

$$C^* = KC^N$$

where K and N are constants.

If the sorption characteristics can be described by a Freundlich sorption isotherm, where C is plotted as a function C^*, the data will be curvilinear. However, the data can be linearized by log transformation by the use of the following equation:

$$\log C^* = \log K + N \log C$$

When log C is plotted against C^*, the result will be linear with a slope of N and an intercept of log K. When the linear intercept (constant N) is equal to 1, the linear isotherm is obtained.

Little is known about the chemical and physical properties of the mineral-bound organic coatings and their potential influence on humic acid sorption to low-carbon sediments. Numerous studies, such as the ones referred to previously, have studied sorption onto well characterized clays and minerals. Natural are a complex matrix of minerals with heterogeneous surfaces on which sorption may occur. The objective of this study to develop a better understanding of the sorptive capacity of natural aquifer materials for humic compounds. This information is necessary to understand the fate and transport of humics in natural aquifer systems so that future sorption studies with metals and other contaminants known to bind to humics can be more clearly understood. This study examined the sorption of humic acids to aquifer materials. Two well characterized standard humics were sorbed to three well characterized aquifer materials. The effects of concentration, pH and ionic strength was studied.

Materials

Sediments:

Aquifer solid materials used consisted of unconsolidated sediments from aquifers at Columbus AFB, Mississippi, Barksdale AFB, Louisiana and Blytheville AFB, Arkansas. The sediments were provided by Dr. T. Stauffer and their collection and treatment are presented in Stauffer (1987) and Libelo (1995) and summarized in Table 1.

Humic Acids:

Two humic acid materials were used in this study, a readily available commercial product, and a well characterized standard material obtained from the International Humic Substances Society (IHSS.) The humic material used for the primary sorption study was Aldrich Humic Acid, Sodium Salt (Aldrich Co., Germany). Aldrich estimated the molecular weight range of their humic material between 5,000 - 500,000 daltons. It has a black, powdery consistency and is quite water soluble. In water, the solutions are deep orange to yellow in appearance dependent on concentrations. (Kim et. al., 1990)

The IHSS humic substance used was Suwannee River Humic Acid Reference 1R101H (100mg.) it was obtained from Dr. P. MacCarthy at Colorado School of Mines. Suwannee River humic acid was isolated from the Suwannee River in Florida which drains the Okefenokee River in southern Georgia. It is a light brown compound, highly water soluble and has a characteristic yellow to pale color in solution depending on concentration. See Table I for physical properties of the humic acids used.

Methods

Rationale for Chosen Methodology and Conditions:

Batch equilibration technique was used to measure sorption of humic acids to aquifer material solids because it 1) allows for the determination of an entire sorption isotherm in a short amount of time, 2.) is commonly used in similar studies of surface soil and sediments, and involves fewer variables than dynamic (column) techniques (Stauffer, 1987).

Sorption Studies at Varying pH:

The humic material used for the preliminary sorption study was the commercial Humic Acid, Sodium Salt obtained form Aldrich Co. A series of 10-fold dilutions from a 1000 mg/L solution were prepared, and a calibration curve was generated as in Davis and Bhatnagar (1995) in order to establish absorbance and concentration parameters. The dissolved humic acid concentration was determined by measuring absorbance at 240 NM using a Carey 3E UV-Visible Spectrophotometer (Varian Instruments) with quartz cuvettes. All isotherm experiments were conducted using 18 OHM Milli-Q water (Millipore Co.) in order to minimize organic and inorganic contaminants. No attempt in this phase of the experiment was made to produce a "synthetic" groundwater by adding inorganic salts or buffers to the Milli-Q water, since each groundwater sample is a product of the particular geologic materials through which it passes and no standardized formula can be made.(Stauffer (1987). The sorption experiments at varying pH were conducted in the absence of buffer solutions. Ionic strength studies did however vary ionic strength buffers and concentrations and was done in a later, independent portion of this study.

Humic Acid Solutions:

Physical characteristics from Table II of the humic materials used in our studies indicates that carbon makes up 49-51% of each sample. The total organic carbon content of natural waters ranges from 1 to 30 mg C/L (Aiken et al. 1985), this experiment utilized concentrations several orders of magnitude above the average in all varying pH isotherm experiments in order to properly study the effects of a range of concentrations of humic acid; from nominal 25-1000 mg.C/L. Each set ranged from nominal 1000 mg/L humic acid (≈350-425 mg C /L) to 25 mg/L (≈11-13 mg C /L). For the IHSS Suwannee humic acid, quantities were limited and so sorption studies were conducted with lower concentrations i.e. ranging from 100 mg/L (≈39-54 mg C/L) to 25 mg/L (≈11-13 mg C/L)

The sample preparation at a given pH on each sediment was conducted in the following manner: The humic acid standards were brought to the appropriate dilution mark with Milli-Q water, stirred at room temperature (22.5^{0}C +/- 1^{0}C) for one hour, purged with N_2 for approximately 20 minutes and then poured into sterile disposable Nalgene® containers for pH determination. The pH values were measured using an Accumet® Model 25 (Fisher-Scientific) instrument equipped with a silver/silver chloride electrode. All pH readings are (+/- 0.1) unit based on manufacturer specifications. The pH of the humic standards was

adjusted by drop-wise additions of concentrated HCl and NaOH (Fisher Co.) respectively. The pH values varied from 1-2 units during the experiment as a result of the natural buffering capacity of the aquifer solid.

A fixed sorbate-solution ratio of 1 gm : 2.5 mL was maintained throughout. The sorbate-solution separations by centrifugation were made as complete as possible to avoid possible sorbate concentration effects. Differences in sorbent density caused the solid-liquid ratio to vary somewhat because the same containers (constant total volume) were used in all isotherm experiments. All data acquired on different sorbents within this work are comparable with regard to solid-liquid ratio.

Aquifer Solids:

The reaction vials were prepared in the following manner: 1 gram (+/- .0005 gm.) of each aquifer solid was weighed into a clear borosilicate, serum bottle (Aldrich Co.) using a Mettler AE 160 analytical balance. Sorption of humic solutions with the serum vial was not significant, ranging from 0.9-2% over a period of 72 hours. The humic acid solution was pipetted onto the sediment. The vials were then capped with butyl rubber stoppers and crimp sealed under an N_2 environment. The vials were placed on a rotary shaker table (Lab-Line, Melrose Park IL) and shaken at 100 RPM for 24 hours. Samples were then centrifuged (Damon/IEC EPR-6000) at 2,500 *g.* for 25 minutes. The supernatant was then used for total organic carbon (TOC) analysis. The total organic carbon present in each sample was quantified using a combustion/non-disperse infrared gas analysis system; (Shimadzu TOC-5000 coupled to a Shimadzu ASI-5000 autosampler, Shimadzu Corp.,Japan) The amount of humic material sorbed onto the sediment was determined by difference from sediment-free samples. The sampling done was batch-process type, with each vial prepared in triplicate,

Ionic Strength Measurements

The procedures employed in this portion of the experiment paralleled those for the non-buffered pH isotherm experiments. However, the study was conducted only on Columbus AFB sediment with Aldrich humic acid because its abundance and, with respect to Barksdale and Blytheville, demonstrated higher sorptivity of the humic material. Two anionic sodium buffers were used ; sodium perchlorate $NaClO_4$ (Aldrich) and Sodium phosphate monohydrate $NaH_2PO_4 \bullet H_2O$ (Fisher). Two concentrations were prepared using each buffer, 0.1M and 0.005 M. The buffer solutions were prepared with Milli-Q water and had a pH upon dilution of 6.7 for the sodium perchlorate and 4.2 for sodium phosphate monohydrate. The solutions were then brought to pH 7(+/-0.1) by drop-wise addition of dilute HCl and or NaOH,. allowed to equilibrate at room temperature for one hour, and then used to prepare the humic acid standards in the same manner as in the pH studies. Slight pH variations occurred when humic acid standards were prepared with the sodium buffer solutions requiring addition of acid or base to obtain the necessary pH. After equilibration, 5 mL of solution aliquots were pipetted into vials containing 2 gm. of aquifer material. The samples were then crimp sealed under an N_2 environment, shaken for 24 hours and

centrifuged for 20 minutes at 20,000 *g*. The supernatant was then analyzed for TOC sorption; determined by difference from sediment-free blanks.

Results and Discussion

The observed trend of TOC sorbed as a function of humic concentration, over all pH ranges was; Columbus > Blytheville > Barksdale. Figures 1-3 of this report represent the isotherms encompassing the nominal pH range 4-7(+/- 0.1 units). The extent of sorption of humic acid substances to three different aquifer material solids decreases with increasing pH. In addition, the sorption constants (K_d) decreased with increasing pH for all sediments; Columbus 2.321 - 0.352, Barksdale 0.323 - 0.079 and Blytheville 0.574 - 0.113. A similar pH dependence was found by Tipping (1981) for the sorption of Esthwaite water humic substances to hydrous iron oxides and by Murphy et al. (1992) for the sorption of various IHSS humic and fulvic acids onto hematite and kaolinite. In contrast to the previously mentioned studies, our study utilized natural aquifer materials. Each type of aquifer solid was extracted from three separate locations from Air Force bases around the southern portion of the United States. Consequently, our samples not only differ in geographical origins but in their physical properties as well. The isotherm plots illustrated in Figures 1-3 of this report show that our low-carbon aquifer sediments behaved similarly to the well classified clay materials of past studies and the results are very much in line with those of the past works mentioned above.

The pH values depicted on the isotherm plots reflect the pH values determined following the 24 hour equilibration time. It should be pointed out that the pH values did not remain at their nominal values, as a result of the sediments natural buffering capacity. For Barksdale and Blytheville sediments, the pH changed 2-3 units from the starting values as compared to a 0.8 - 1 pH unit change for the Columbus sediment. The final pH value range after equilibration for Barksdale was 6.79 - 8.03, for Blytheville and Columbus sediments the ranges were 6.57- 7.98 and 4.83 - 7.79 respectively. This buffering capacity most likely is the result of carbonate, or phosphate species on the aquifer sediments most likely from the clay minerals.

The physical properties of the aquifer materials (Table 1), may correlate with the observed trend seen with pH variation. Columbus AFB sediment sorbed the highest amount of TOC, (29-80% at pH 4) of the humic material added. Columbus sediment has a considerable surface area (5.74 m^2/gm) and surface iron composition (0.914 g/kg). In addition, Columbus sediments are about 70.8% sand. Conversely, Columbus is made up of a significant percentage of silt and clay; 17.50% and 11.68% respectively. This suggests that it is the sites on the silt and clay that are most likely involved in the sorption of the humic material. They contain the surface hydroxyl sites involved in surface metal oxide interaction and ligand exchange mechanisms. This is further illustrated by comparison of the physical properties of Blytheville sediment sorption isotherms. Although Blytheville has the largest surface area, (9.32 m^2/g), and dithionite iron composition, (1.64 g/kg) of the three sediments studied, it is comprised of a substantial percentage of sand, 95.6% and subsequently a rather small percentages of silt and clay; 3.06% and

1.34% respectively. Sand does not provide sites where sorption may occur, consequently both Blytheville and Barksdale aquifer sediments sorbed only 2-30% (at pH 4) of the humic acid in solution at equilibrium less than Columbus. This can be attributed to their low percentages of clay and silt compositions, resulting in fewer sites for sorption.

Several explanations regarding the effect of pH variation on humic sorption could assist in explaining the observed trend. Primarily, the sorption of humic substances increases with decreasing pH in response to positive charge development on the sorbents (Tipping, 1981). Protons may neutralize the negative charges on humic anionic groups. not involved in the adsorptive interaction. and decrease electrostatic repulsion when the HS concentrates at the interface. (Tipping, 1981; Tipping and Cooke; 1982). In addition this observation is consistent with a ligand exchange mechanism previously suggested by Tipping, 1981; Tipping and Cooke, 1982; Davis, 1982; and Murphy et al., 1992. Sorption of humic substances by ligand exchange is believed to occur in the following sequence (Sposito, 1984):

(1) $SOH + H^+ \Longleftrightarrow SOH_2^+$

(2) $SOH_2^+ + HuC(O)O^- \Longleftrightarrow SOH_2^+\text{-}OC(O)\text{-}Hu$

(3) $SOH_2^+\text{-}O\text{-}C(O)\text{-}Hu \Longleftrightarrow SO\text{-}C(O)\text{-}Hu + H_2O$

SOH represents the surface hydroxyl group on the sorbent. and $HuC(O)O^-$ represents the humic carboxyl group. Equation (1) represents the protonation step believed to render the surface hydroxyl group more exchangeable (Sposito, 1984). The protonation step is responsible in part, for the pH dependence of the fractional sorption. Equation (2) represents the formation of the humic carboxyl group, an outer-sphere surface complex with a protonated hydroxyl group. Equation (3) represents ligand exchange (OH_2 for Hu-COO-) that is postulated to occur.

The second humic acid was sorbed onto these aquifer sediments to determine if a different type of humic acid would behave differently than the commercial non-purified Aldrich humic material. The results of Murphy et al. (1992) point out that the sorption of humic substances was directly proportional to the aromaticity and inversely proportional to the polarity as approximated by the elemental O/C ratio. Table 2 compares some of the physical properties of the Aldrich and Suwannee humic acid. The percent aromaticity and O/C ratios of Aldrich and Suwannee River humics are comparable. Accordingly, sorption isotherms (Figures 7 and 8) show that sorptions are quite comparable. The isotherms satisfactorily depict the initial, linear portion of the isotherm. It is clear that these isotherms are in reasonably good agreement with the conclusions of Murphy et.al., (1990) that O/C ratios and percent aromaticity are the key factors which govern humic acid sorption. Table 4 illustrates the results of the statistical treatments for the sorption of Aldrich and Suwannee humic acid on to Barksdale and Blytheville sediment.

The data for the isotherm plots of Figures 1-3, 7 and 8 were treated by statistical methods to determine the best fit parameters to our data values. Tipping (1981) and Murphy et. al. (1992) determined that Langmuir fits were the most statistically favorable in the treatment of their data. However, as pointed out by Stauffer (1987), the choice of statistical methods applied to humic acid sorption is often varied and no mathematical treatment of the data is uniformly accepted as better than

another. In this study, the sorption data were fit to Linear, Langmuir and Freundlich isotherms to determine the best statistical fit. Figures 4-6 illustrate the different isotherm fits for Aldrich humic acid at nominal pH = 4. Linear and Freundlich isotherm plots fit our data better than the Langmuir fits. This may be attributed to the fact that our natural sediment surfaces are heterogeneous rather than well defined homogeneous clays and minerals used by Tipping (1981) and Murphy et al.(1990). The linear K_d and Freundlich log K (intercept) and N (slope) are shown in Tables 3, 4 and 5.

Two different ionic strength studies were conducted using two compounds with distinctively different chemical properties and known effects on humic acid sorption. The compounds used were sodium perchlorate ($NaClO_4$ and sodium phosphate monobasic ($NaH_2PO_4 \bullet H2O$). The results of the ionic strength isotherms were consistent with those found by Tipping(1981); Tipping and Cooke (1982); Tsutsuki and Kuwatsuka (1981) and Murphy et. Al. (1992). Figure 9 of this report illustrates the effects of each buffer at high and low ionic strength; 0.1 M and 0.005 M respectively. As pointed out in the introduction, ionic strength and pH variation have an effect on the functional groups of the humic acids. These functional group changes thereby alter the sorptive capacity of the humic acid onto the aquifer material solid. The aquifer material used in this study was Columbus AFB solids. This particular aquifer material was selected because in comparison to Barksdale and Blytheville solids, Columbus sediments supported noticeable sorption and a complete isotherm could be followed readily within our detection limits.

Sodium perchlorate is a relatively strong oxidizing agent and has the capacity to maintain it's ClO_4^- ion in solution. The results of the sodium perchlorate buffered humic acid isotherms appear to support the conclusions of the Murphy et al., (1992) that humic acids seem to adopt a more linear or open configuration at low ionic strength, and on a mole carbon basis, occupy a greater amount of mineral surface than humic acids at high ionic strength. In addition, our results supported the findings of Tsutsuki and Kuwatsuka (1984) which stated that the humic acid molecule is fully expanded at low ionic strength at pH 7.5, but folds onto itself considerably with increasing ionic strength. Due to the fact that the humic acid tends change its structural configuration, it subsequently may have less exposed ionizable carboxyl and hydroxyl sites available for sorption onto the sediment. For reference purposes, the isotherm of a non-buffered humic acid solution is plotted in Figure 9, and as is evident, the amount of TOC sorbed is quite a bit less ≈35-40% then in the 0.1 M perchlorate-buffered isotherm.

Tipping (1981) pointed out that in some instances, the presence of certain anion species like phosphates (PO_4^-) can effect the sorption of humic acids onto iron oxide materials because they directly compete with the humic acid for the surface hydroxyl site on the sediment. As is depicted in Figure 9, the presence of the phosphate ion greatly reduces the TOC sorbed onto the natural aquifer material. Comparison of the TOC sorbed between 0.1 M and 0.005 M sodium phosphate demonstrate that the differences in concentration by a factor of 20 are not dramatic. This is most likely due to the fact that the phosphate loading capacity on the sediments has been achieved and excess phosphate may itself bind to a portion of the humic acid or remain in the colloidal suspension with the remainder of the humic acid.

The results are however notable when one compares the phosphate buffered isotherm to the non-buffered isotherm. The amount of TOC sorbed in the presence of phosphate ions is $\approx$ 26-48 % less than that seen in the phosphate-free reference isotherm. Most importantly, it is clear from this low TOC uptake that a competition for the aquifer site exists between the PO_4^- and the larger, more bulky humic material. This is the first study of whole sediments not specific minerals with phosphate species. This study not only gives a more accurate understanding of the interaction of natural aquifer materials with phosphates it gives a much better understanding of overall sorptive capacities of these natural materials.

Conclusions and Future Work

The results of this experiment will allow new insight into the sorption capacity of natural aquifer materials for humic acids under varying conditions of pH, concentration and buffer species. This information will be important for further sorption studies involving the use of additional sorptive species such as metals and environmental pollutants. Preliminary studies regarding the desorption of the humic materials from the sediments were conducted but need to be explored further. It will be important in future studies incorporate more humic and fulvic acids than the ones used here. In addition, more spectroscopic techniques such as atomic adsorption and fluorescence spectroscopy should be employed to further examine the role of Fe(II) and Fe(III) species whose presence is important for numerous environmental systems.

Table 1. Chemical and Physical Characteristics of Aquifer Materials[1].

Aquifer	Total Organic Carbon % (sd) n = 5	Surface Area m^2/g (sd) n = 5	Dithionate Iron g/kg (sd) n = 5	% Sand	% Silt	% Clay
Barksdale	0.0338 (0.0038)	0.88 (0.22)	0.112 (0.04)	96.1	1.94	1.97
Blytheville	0.0676 (0.0047)	9.32 (0.05)	1.64 (0.317)	95.6	3.06	1.34
Columbus	0.0596 (0.0062)	5.78 (0.049)	0.914 (0.067)	70.8	17.50	11.68

1. Libelo (1995)

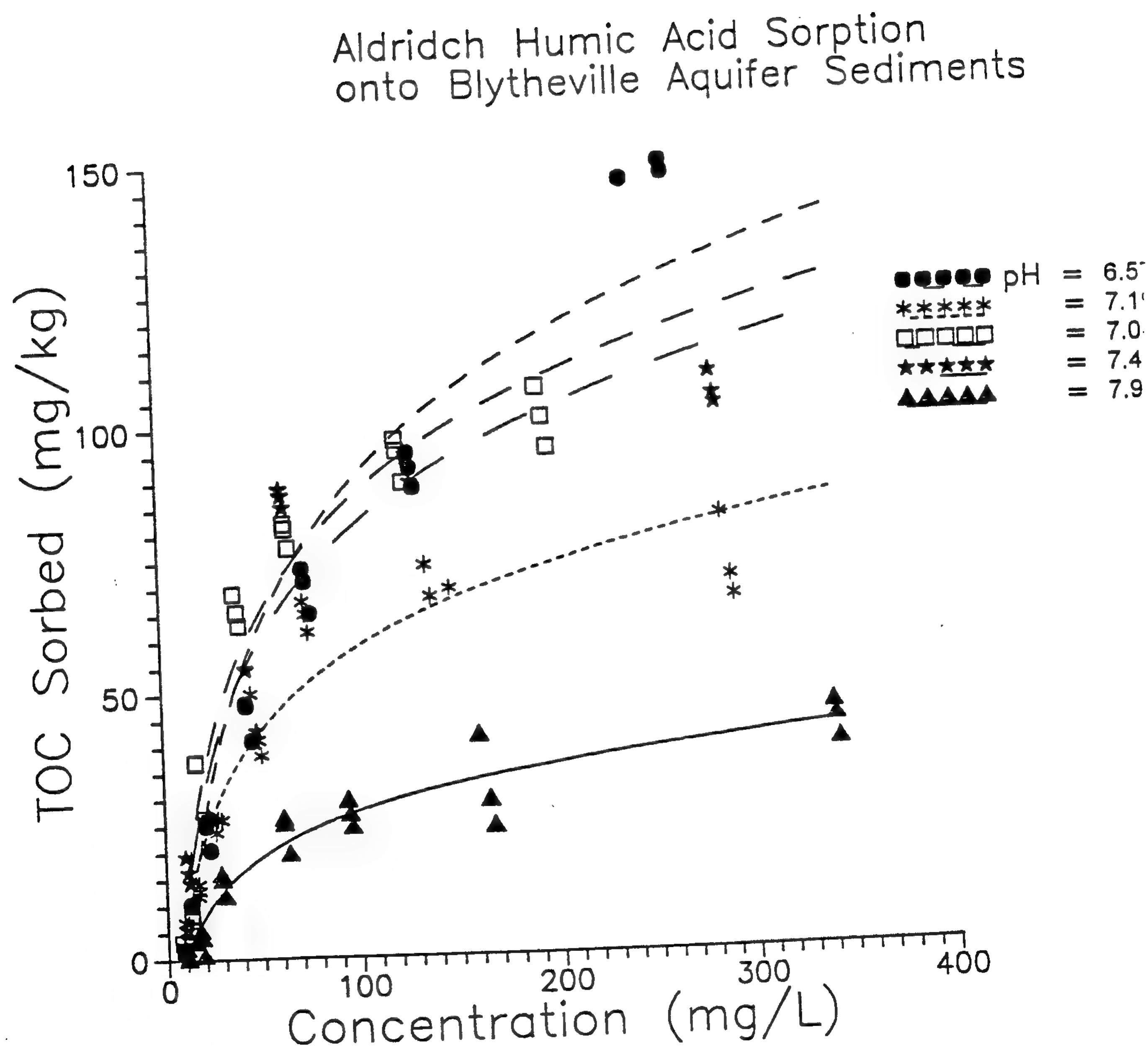
Table 2.

Physical properties of the humic acids.

Sample	Elemental Composition C	H	N	S	O	Other	Atomic Ratios H/C	O/C	Aromatic C (110-165)	Aliphatic C (0-90)	Carboxyl C (165-190) PPM
Aldrich Humic Acid[a]	41.72	4.37	0.25	1.90	36.93	14.83	1.03	0.51	40	41	14
Suwannee River Humic[b,c] (1R1O1H)	50.50	3.35	1.18	0.55	44.42		1.26	0.66	37	28	19

a: Kim et al., 1990
b: Thurman and Malcolm, 1981
c USGS Technical Report, 1989

Figure 2.



Aldridch Humic Acid Sorption
onto Blytheville Aquifer Sediments
TOC Sorbed (mg/kg)
Concentration (mg/L)
150
100
50
0
0
100
200
300
400
pH = 6.5
= 7.1
= 7.0
= 7.4
= 7.9

Figure 3.

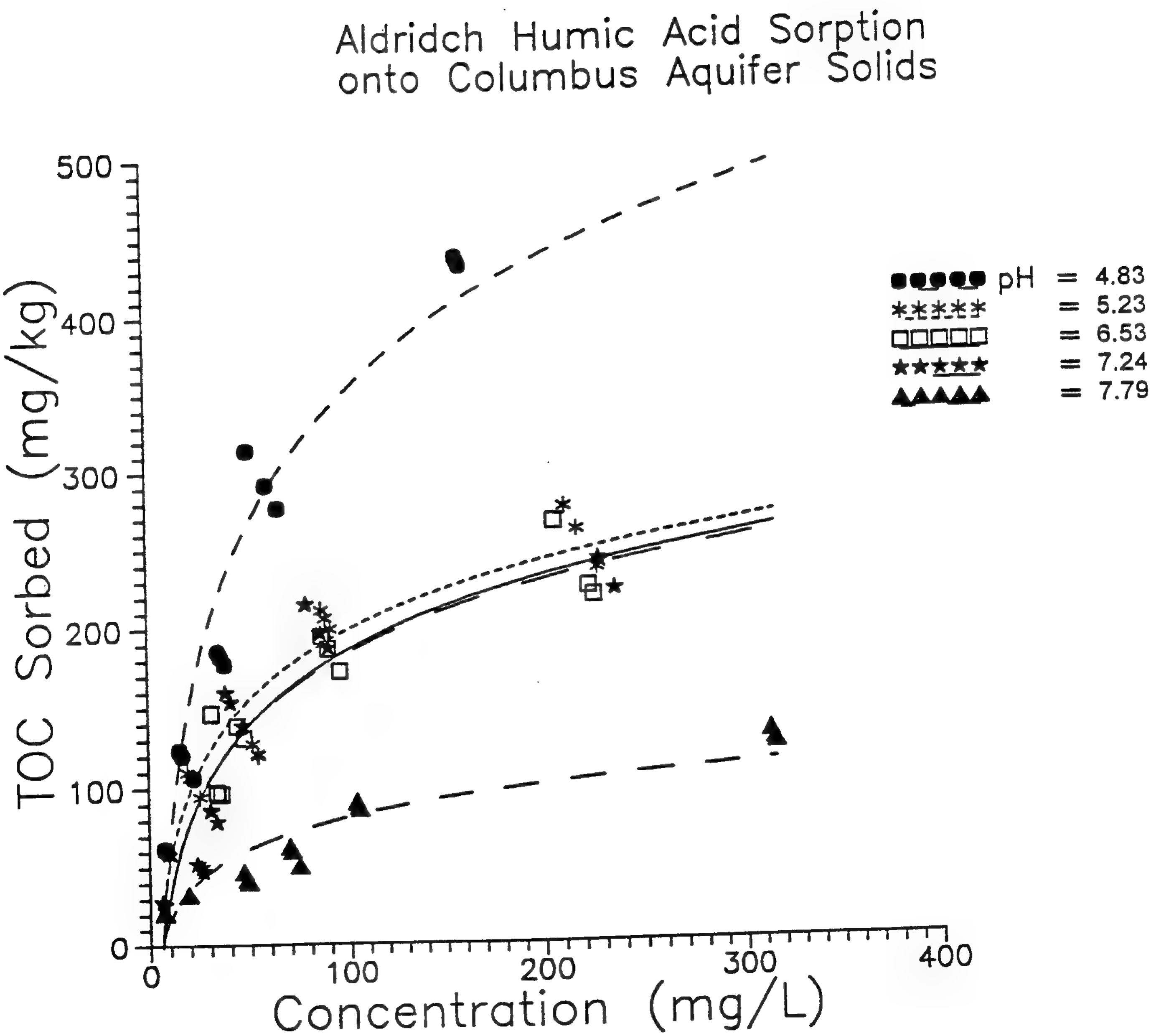

Aldridch Humic Acid Sorption
onto Columbus Aquifer Solids
TOC Sorbed (mg/kg)
0
100
200
300
400
500
Concentration (mg/L)
0
100
200
300
400
pH = 4.83
= 5.23
= 6.53
= 7.24
= 7.79

Figure 4.

Aldridch Humic Acid Sorption onto Columbus Sediments pH = 4.83

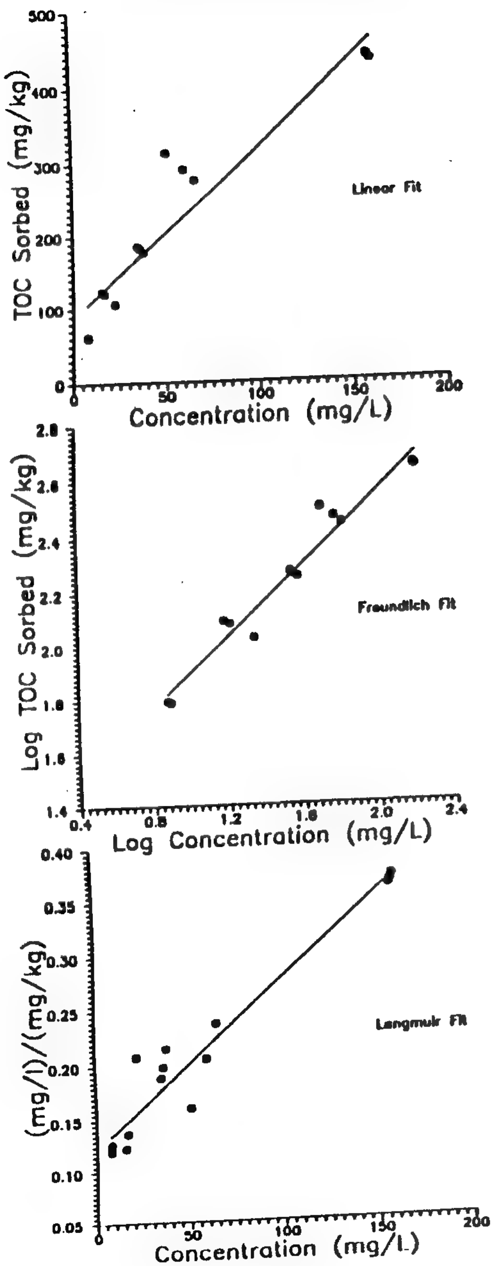


Figure 5.

Aldridch Humic Acid Sorption onto Barksdale Sediments pH = 6.792

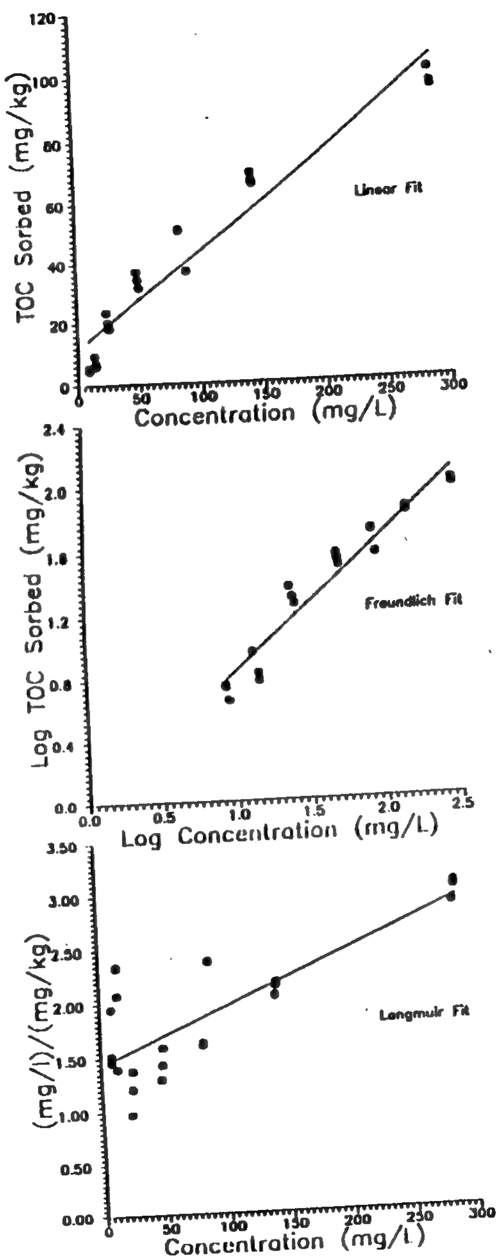


Figure 6.

Aldridch Humic Acid Sorption onto Blytheville Sediments pH = 6.575

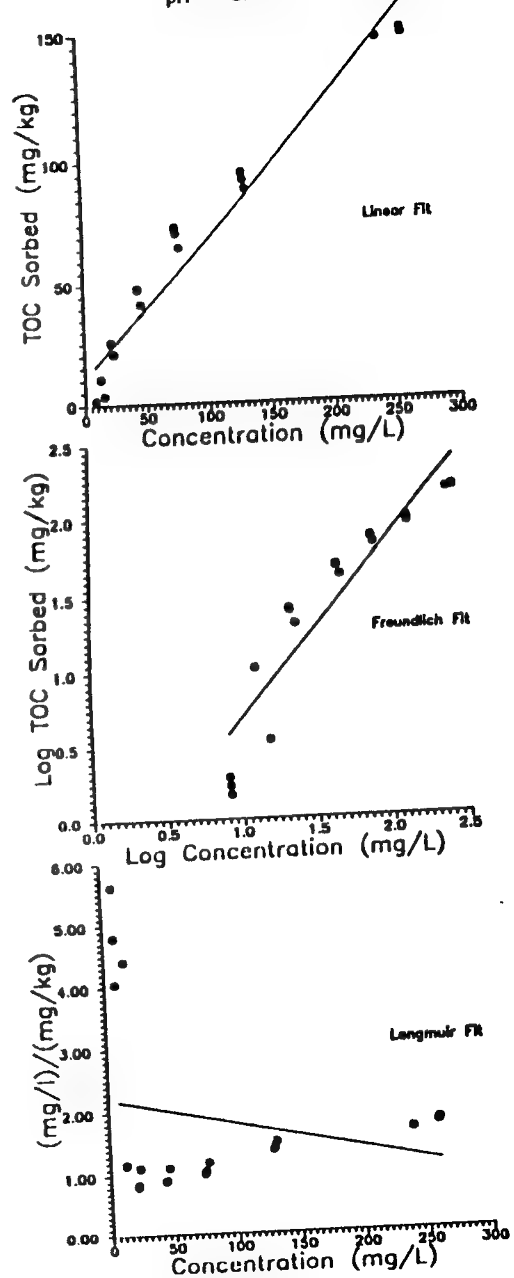


Figure 7.

Sorption of two humic acids on aquifer material solids. (Initial pH = 4)

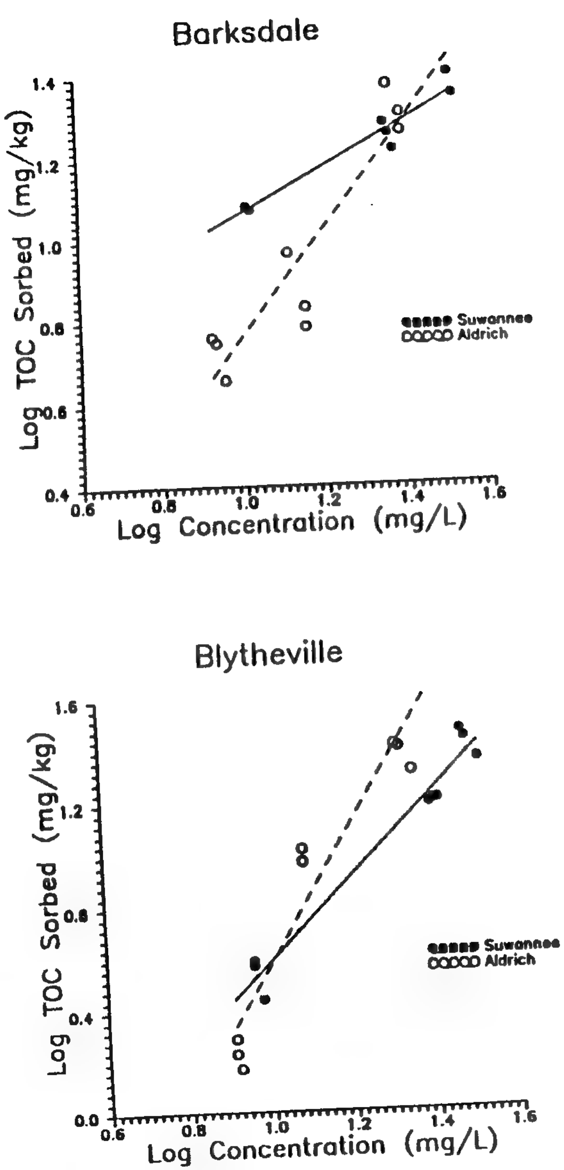


Figure 8.

(Initial pH = 7)

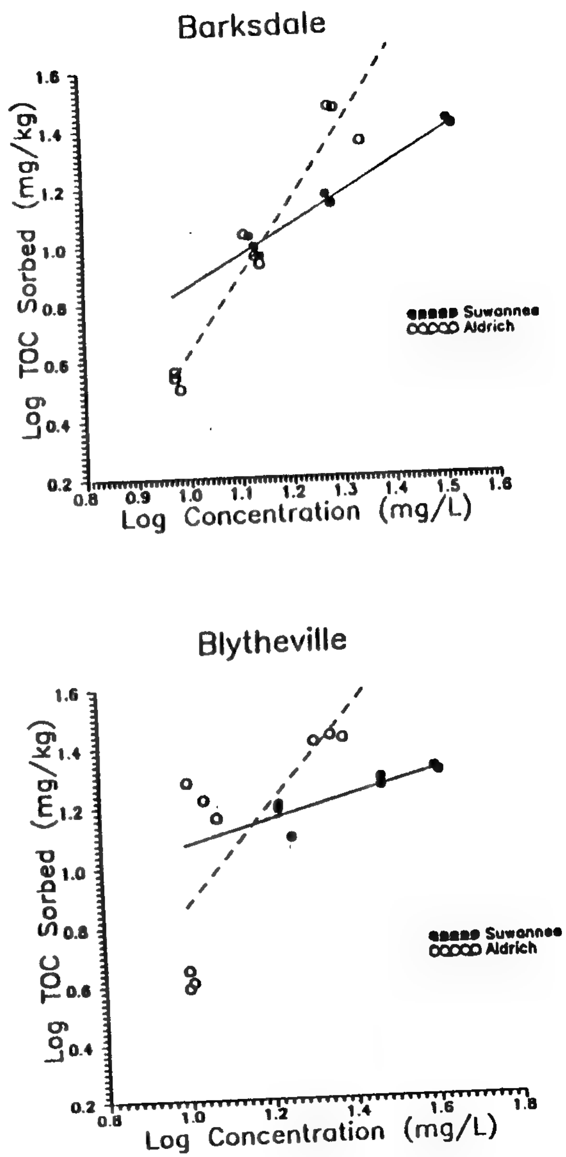


Table 4

Isotherm results of Aldrich and Suwannee River humic acid sorption on aquifer solids.

Sediment	Humic Acid	pH[a]	Linear Fit K_d[b]	Freundlich Fit log K	Freundlich Fit N
Barksdale	Suwannee	4	0.480 (0.048)	6.7E-16 (1.5E-16)	1.0 (2.4E-16)
	Suwannee	7	0.807 (0.035)	-0.216 (0.028)	1.067 (0.060)
	Aldrich	4	0.996 (0.149)	-0.529 (.112)	1.298 (.2032)
	Aldrich	7	2.065 (0.323)	-1.985 (0.102)	2.608 (0.251)
Blytheville	Suwannee	4	0.936 (0.119)	-1.007 (0.078)	1.588 (0.113)
	Suwannee	7	0.236 (0.043)	-0.687 (0.037)	0.384 (0.079)
	Aldrich	4	1.554 (0.152)	-2.003 (0.174)	2.581 (0.331)
	Aldrich	7	1.307 (0.326)	-0.729 (0.260)	1.598 (0.544)

a: Initial pH of humic acid solution.
b: Linear least-squares regression of data.
( ) standard error of coefficient

Table 3.

Isotherm results of Aldrich humic acid sorption on aquifer solids.

Sediment	pH[a]	Linear Fit K_d[b]	Freundlich Fit log K	Freundlich Fit N
Columbus	4.83	2.321 (0.213)	1.251 (0.062)	0.649 (0.035)
	5.23	0.884 (0.104)	1.351 (0.06)	0.462 (0.033)
	6.53	0.641 (0.105)	1.426 (0.067)	0.411 (.061)
	7.24	0.846 (0.141)	0.918 (0.127)	0.664 (0.064)
	7.79	0.352 (0.028)	0.908 (0.050)	0.476 (0.016)
Barksdale	6.79	0.323 (0.021)	0.0041 (0.111)	0.846 (0.048)
	6.795	0.175 (0.021)	0.075 (0.445)	0.677 (0.174)
	6.908	0.164 (0.03)	0.149 (0.255)	0.726 (0.097)
	7.563	0.162 (0.029)	0.181 (0.224)	0.726 (0.096)
	8.03	0.079 (0.012)	-0.380 (0.257)	0.798 (0.014)
Blytheville	6.57	0.574 (0.031)	0.50 (0.247)	1.191 (0.107)
	7.19	0.222 (0.040)	0.231 (0.155)	0.747 (0.074)
	7.04	0.492 (0.069)	-0.12 (0.27)	1.035 (0.121)
	7.43	0.338 (0.055)	0.182 (0.231)	0.841 (0.101)
	7.98	0.113 (0.016)	-0.801 (0.349)	1.063 (0.151)

a: pH of supernatant after equillibration
b: Linear least-squares regression of data
( ) standard error of coefficient

Figure 9.

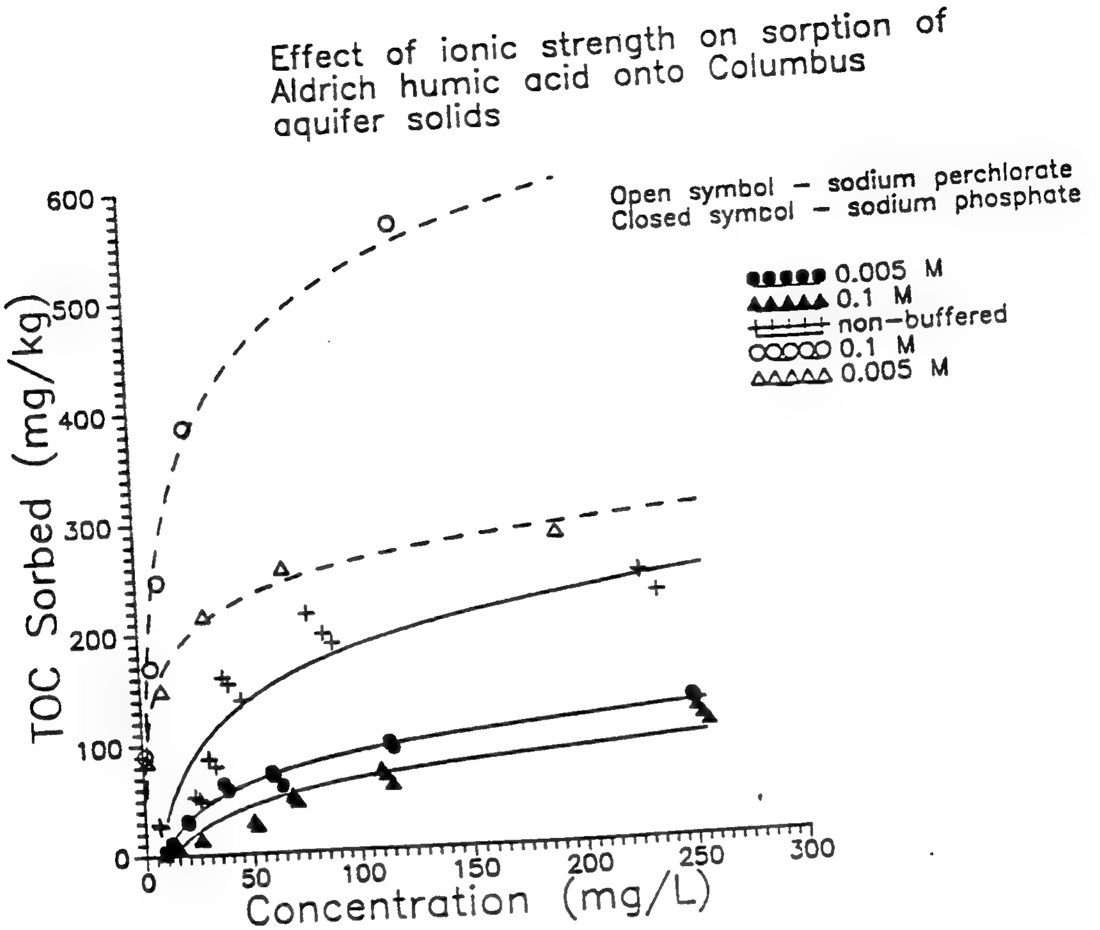


Table 5.

Isotherm statistical sorption results of buffered humic acid solutions onto Columbus AFB aquifer solids.

Buffer	Ionic Strength	Linear Fit K_d	Freundlich Fit log K	Freundlich Fit N
Sodium Phosphate	0.005 M	0.461 (0.048)	-0.01 (0.238)	0.969 (0.105)
	0.1 M	0.439 (0.019)	-0.580 (0.162)	1.16 (0.074)
Sodium Perchlorate	0.005 M	0.819 (0.317)	1.917 (0.357)	0.258 (0.027)
	0.1 M	3.248 (0.904)	1.937 (0.059)	0.415 (0.041)
Non-Buffered		1.273 (0.15)	0.918 (0.128)	0.664 (0.064)

SEROTONERGIC INVOLVEMENT IN PHOTIC-INDUCED PHASE ADVANCES OF HAMSTER CIRCADIAN WHEEL RUNNING ACTIVITY

Anne Riberdy
Department of Psychology
Boston University
Boston, MA 02215

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC
and
Biological Rhythms and Integrated Neuroscience (BRAIN) Institute
Armstrong Laboratory

August 1995

RECEIVED SEP 2 0 1995

SEROTONERGIC INVOLVEMENT IN PHOTIC-INDUCED PHASE ADVANCES OF HAMSTER CIRCADIAN WHEEL RUNNING ACTIVITY

Anne F. Riberdy
Department of Psychology
Boston University

Abstract

A hypothesis is presented which proposes that serotonin $(5\text{-HT})_{1B}$ autoreceptors are located on retinal afferents which project to the suprachiasmatic nucleus (SCN). It is theorized that 5-HT binding at retinal afferents exerts an inhibitory effect on glutamate at the level of the synapse between the retinohypothalamic tract (RHT) and the SCN. Receptor autoradiography as well as behavioral experiments were carried out in order to substantiate this claim. Results suggest that although serotonin does appear to modulate glutamate release at the level of the SCN, the 5-HT receptor subtype which mediates this mechanism is unclear.

Introduction

The suprachiasmatic nucleus (SCN), located immediately dorsal to the optic chiasm in the anterior hypothalamus, engenders and maintains the circadian rhythms of physiological and behavioral processes in mammals. This endogenous timekeeping mechanism is responsible for synchronization or entertainment of the SCN to the light/dark cycle predominantly through photic input relayed from the retinohypothalamic tract (RHT). The RHT consists of sensory afferents originating in the retina which project bilaterally to the SCN[1,2]. Additional, however less substantial input to the SCN mediating entrainment originates in the intergeniculate leaflet (IGL), which also receives retinal afferents.

Excitatory amino acids (EAAs) such as glutamate have been found to act as neurotransmitters at the synapse of the RHT, where photic information is relayed to the SCN[3,4]. Consequently, it is likely that EAA neurotransmission mediates signaling between the RHT and SCN. Indeed, antagonists of EAAs lock synaptic transmission between the RHT and SCN *in vivo* and *in vitro*[5,6]. In addition, EAA antagonists block functions governed by the RHT, such as the SCN's response to optic nerve stimulation, light induced phase shifts in circadian activity rhythm and light induced Fos expression in the SCN[5]. All of these actions are presumably a function of EAA transduction of photic stimuli.

The SCN also receives substantial serotonergic innervation from the midbrain raphe nuclei[7]. Termination of raphe fibers in the SCN coincides anatomically with RHT terminals, suggesting an anatomical basis for interaction between these systems[7,8]. Ample evidence exists which implies this serotonergic innervation is able to modulate the SCN's response to photic stimulation through the inhibition of glutamate release at RHT terminals[7,9]. Serotonergic agonists act in a manner similar to that of EAA antagonists. Local administration of quipazine, for example, results in a diminished response of the SCN to photic stimulation[9]. In addition, serotonin has been reported to attenuate the response of the SCN to excitatory optic nerve input[10]. Such evidence suggests that serotonin may exert a modulatory effect on EAAs such as glutamate. Presumably, then, 5-HT binding at retinal afferents may mediate glutamate release at the synapse between the RHT and SCN.

It has been suggested that 5-HT_{1B} receptors are located on axon terminals, and might possess an addressing signal which allows for transport of the receptor protein towards nerve ending[11]. In addition, 5-HT_{1B} autoreceptors have been shown to inhibit neurotransmitter release from nerve terminals presynaptically. Although no direct evidence exists that 5-HT_{1B} receptors are located on axon terminals in the SCN, this does appear to be the case in the superior colliculus (SC). 5-HT_{1B} binding in the SC has been shown to

decrease substantially subsequent to ocular enucleation, whereby retinal afferents along with their terminals would be destroyed. The superficial grey layer of the SC, like the SCN, receives both retinal and serotonergic innervation. Receptor autoradiography using [^{125}I]iyodocyanopindolol ([^{125}I]ICYP) has demonstrated that the SCN contains 5-HT$_{1B}$ autoreceptors[12]. These lines of evidence suggest that the 5-HT$_{1B}$ receptor may play a role in the mediation of glutamate release from retinal afferents in the SCN. Thus, should 5-HT$_{1B}$ binding in the SCN be reduced following enucleation, support for presynaptic 5-HT$_{1B}$ modulation of photic input to the SCN via inhibition of the glutamatergic system would be provided.

The purpose of this series of experiments is to: 1) determine whether presynaptic 5-HT$_{1B}$ receptors are located on retinal terminals in the SCN; and 2) determine whether 5-HT, acting via the 5-HT$_{1B}$ receptor modulates photic input to the SCN via the inhibition of glutamate release. Based on previous studies in which ocular enucleations resulted in decreased [^{125}I]ICYP binding in the SC[13], we may find that degeneration of retinal afferents will result in diminished 5-HT$_{1B}$ binding in the SCN as well. In addition, in order to assess the potential role of the 5-HT$_{1B}$ receptor in the modulation of photic input to the circadian pacemaker, we will test whether administration of a selective 5-HT$_{1B}$ agonist will inhibit light induced phase shifts.

Experiment 1

Effects of local CP-93,129 and TFMPP infusion on photic-induced phase shifts of circadian wheel running activity

Introduction

Shifts in animals' circadian wheel running activity will be assessed following local administration of the serotonergic agonist TFMPP and the selective 5-HT$_{1B}$ agonist CP-93,129 with light stimulation in order to elucidate whether 5-HT$_{1B}$ autoreceptor agonists inhibit photic input to the SCN. Should this compound block the phase shifting effects of light stimulation, results will provide some evidence for serotonergic involvement in light-induced phase shifts, presumably through the inhibition from glutamate release from retinal afferents.

Methods and Materials

Subjects Adult male Syrian hamsters *Mesocricetus auratus* (Charles River, MA) weighing approximately 150g at the time of experiment. Animals were originally housed six to a cage

in a vivarium maintained on a 14:10 light-dark (LD) cycle (lights on at 20:00 hr) for at least 14 days prior to surgery. Food and water were supplied ad libitum.

Surgical Procedure Animals were anesthetized with a cocktail of ketaset (125 mg/kg), xylazine (20 mg/kg) and acepromazine maleate (2 mg/kg) injected intraperitoneally (i.p.). Cannula guides (26 gauge) containing 33-gauge stylets (Plastic Products, Roanoke, VA) were implanted to a depth of 2.9 mm ventral to the dura. Cannula guides were oriented such that the infusion cannula used for drug delivery would target the SCN. Stereotaxic coordinates were 1.0 mm anterior to bregma and 1.5 mm lateral to midline at a vertical angle of 10∞ (upper incisor bar at 0°). Cannulae were then fixed to the skull with fine screws and dental cement. The wound was closed with suture material and the animals were allowed a recovery period of at least 7 days.

Activity Rhythms Immediately following surgery, animals were placed in individual cages containing 7-in running wheels (Columbus Instruments, Columbus, OH) and maintained in constant darkness (DD). Wheel running activity was continuously monitored using a Zenith 248 computer running Dataquest III data acquisition software (supplied by Mini Mitter, Sunriver, OR). Actograms were generated and analyzed using Circadia software (Behavioral Cybernetics, Cambridge, MA).

The onset of wheel running activity, defined as circadian time (CT) 12 was used as a reference point for the timing of injections and photic stimulation. Activity onset was defined as the first 6 minute interval that 1) coincided with a period of activity that exceeded 10% of the maximum rate for the day, 2) preceded by no less than 4 hrs of inactivity, and 3) followed by a period of at least 1 hr of activity. The period of free running activity, designated as τ, was calculated as the average time between first activity onset onsets over 5 days prior to experimental manipulations. Activity onsets for the day of the experiment were extrapolated from the least squares regression line fitted through activity onsets over the preceding 5 days.

Drugs CP-93,129 (Pfizer, Inc.) is a serotonergic agonist selective for the 5-HT_{1B} receptors. It does not cross the blood brain barrier.

TFMPP (N-(3-Trifluoromethylphenyl)piperazine) (Research Biochemicals, Inc.) is a serotonergic agonist with an affinity for 5-HT_{1B} receptors.

Drug injection procedure CP-93,129 and TFMPP were dissolved in 0.9% saline to make a 1mM solution. Sterile saline served as vehicle. CP-93,129 or TFMPP (0.3ml of 1mM solution) was administered locally using a 33-gauge infusion cannula attached to a 1-ml Hamilton syringe. The infusion cannula extended 4.4 mm beyond the tip of the cannula guide to target the dorsal border of the SCN.

Light induced phase shifts After at least 10 days in constant darkness (DD), groups of hamsters received injections of either drug or vehicle 10 min prior to CT 19, under dim red light (< 1 lux). Subjects were immediately returned to DD in the home cages. Animals received light stimulation (10 min of 20 lux) 10 min subsequent to injections. The light stimulation chamber has been previously described[14]. Light intensity was measured using a Tektronix J16 digital photometer with J6511 illuminance probe. Animals were returned to DD in their respective cages immediately following light stimulation.

Phase shift quantification Animals remained in DD and wheel activity was monitored for 8 to 10 days after experimental manipulations. Shifts in phase activity were determined by calculating the difference between the intercepts of the 1) forward extrapolation of the least squares regression line through the 5 days immediately preceding stimulation, and 2) backward extrapolation of the least squares regression line through days 3 to 8 following stimulation, or when a stable period was re-established. After data collection, animals were killed and the brains removed and fixed in formalin. The location of the injection for each animal was verified histologically through examination of 100μm thick sections cut through the injection site.

Results

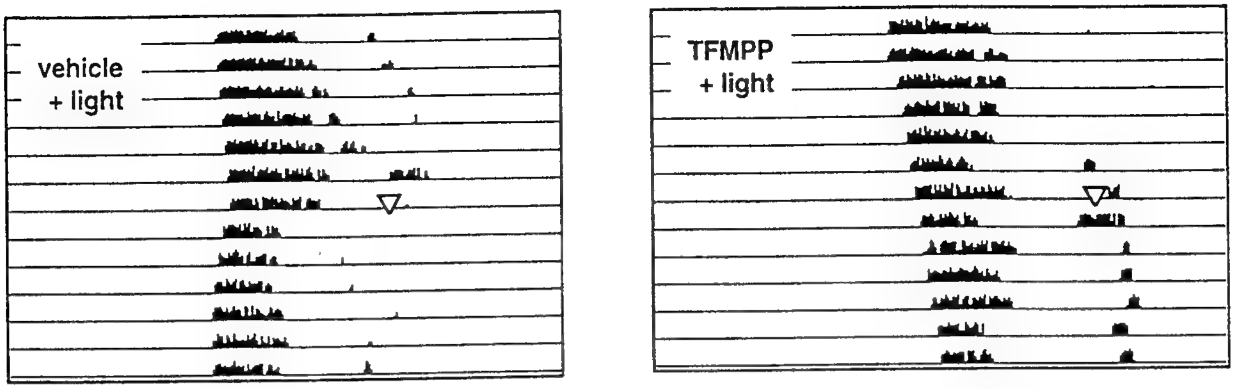
Local administration of TFMPP 10 min prior to light stimulation at CT 19 attenuated light induced phase advances by approximately 90% (0.07 ± 0.06 hrs) compared to control animals (0.71 ± 0.17 hrs) in free running activity. Local administration of CP 93,129, however, did not effectively block the phase advance induced by light stimulation (0.66 ± 0.12 hrs), and resulted in a shift comparable to that of vehicle. Representative actograms displaying wheel running activity for each group can be found in Fig 1A.

Discussion

The finding that local administration of CP-93,129 did not attenuate photic induced phase advances did not support our hypothesis that the 5-HT_{1B} receptor modulates photic input to the SCN. This would also suggest, according to our original hypothesis, that presynaptic 1B autoreceptors do not modulate excitatory amino acid release at the level of the SCN. Although evidence for such a conclusion seems apparent, such an interaction between these systems should not be ruled out. Further investigation should elucidate whether or not our dose regimen, and the time course allowed for CP-93,129 to take effect was efficacious.

The finding that local administration of TFMPP resulted in almost complete inhibition of photic induced phase shifts, while CP-93,129 did not, suggests that the inhibitory effects of TFMPP might be due to action at a serotonin receptor besides the 1B subtype. In addition, to

Fig. 1A. Effects of local TFMPP, CP 93,129 and vehicle infusion on the circadian rhythm of wheel-running activity on hamsters.



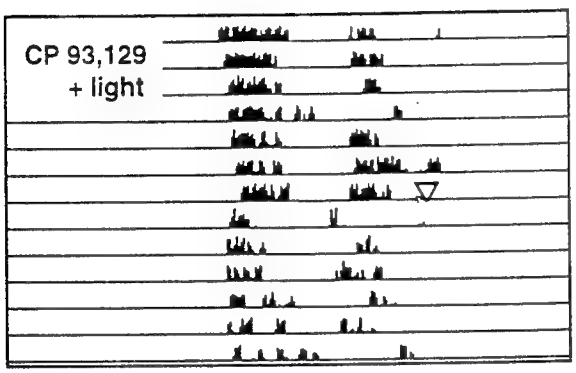


Fig. 1A Representative actograms displaying the circadian rhythmicity of hamster wheel running activity. TFMPP, CP 93,129 or vehicle was infused into the hamster SCN 10 min prior to light stimulation at CT19 (denoted by ▽). Infusion of TFMPP (0.3 µl of 1mM solution) effectively blocked 90% of the phase advance resulting from photic stimulation. CP 93,129 (0.3 µl of 1mM solution) did not block light-induced phase advance and was no different from vehicle.

5-HT_{1B} TFMPP has been reported to have affinity for a number serotonergic receptor subtypes, including the 1A[15,16], 1C[17], 1E[18,19], 2A[20], and 2B[19,20,21]; this list is not exhaustive. Specific serotonergic antagonists should be administered in conjunction with TFMPP to demonstrate the specific receptor subtype responsible for TFMPP's inhibition of phase advances. Furthermore, a selective antagonist for the 5-HT_{1B} receptor needs to be developed and tested before any firm conclusions can be drawn from the data obtained thus far.

Experiment 2

Attenuation of Light-Induced Phase Advances by TFMPP

Introduction

In order to elucidate whether serotonergic agents inhibit photic input to the SCN, shifts in animals' circadian wheel running activity were assessed following the systemic administration of TFMPP. TFMPP was administered 30 min prior to light stimulation in a series of doses ranging from 0.0 to 5.0 mg/kg. In light of the results obtained in Experiment 1

of this report, we hypothesized that the dose of TFMPP should be inversely correlated with the magnitude of phase advances induced by light stimulation at CT 19. That is, TFMPP should result in the attenuation of light-induced phase shifts in a dose dependent manner.

Materials and Methods

Subjects Adult male Syrian Hamsters Mesocricetus auratus (Charles River, MA) weighing approximately 150g at the time of experiment. Animals were originally housed six to a cage in a vivarium maintained on a 14:10-h light-dark (LD) cycle (lights on at 20:00 hr) for at least 14 days prior to transfer to constant darkness (DD) in individual cages containing 14-cm running wheels. Food and water were supplied ad libitum.

Activity rhythms Wheel running activity was monitored and shifts in animals' circadian wheel running activity were quantified in accordance with the *Methods* section of Experiment 1 of this report.

Light induced phase shifts After 10 days under DD, groups of hamsters received i.p. injections of either vehicle, 0.125, 0.25, 0.5, 1.0, or 5.0 mg/kg TFMPP, 30 min prior to CT 19 under dim red illumination (< 1 lux), then were immediately returned to darkness. Hamsters received light stimulation 30 min following injections, at CT 19 (10 min of 20 lux). Animals were returned to DD in their respective cages immediately following light stimulation.

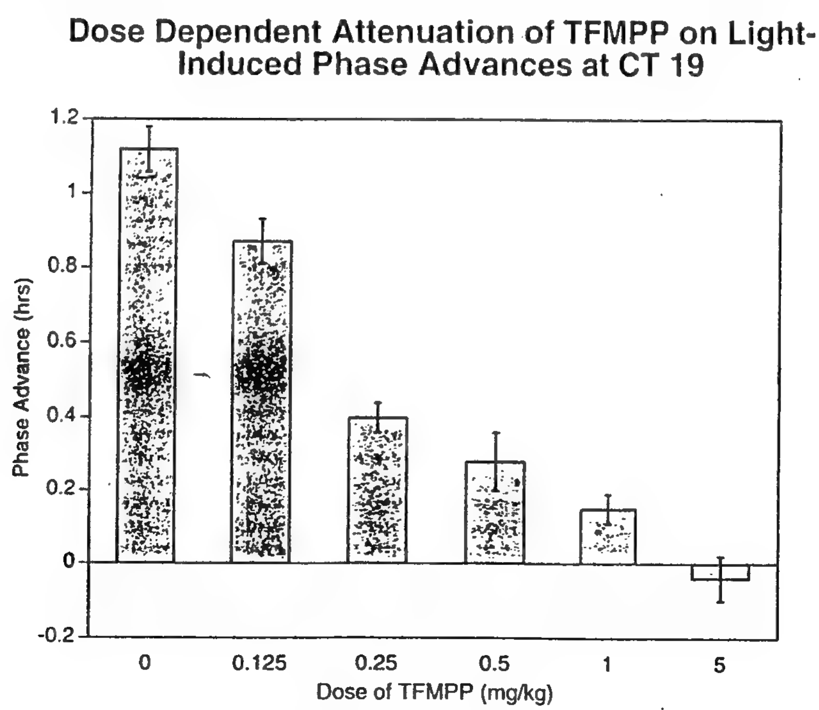


Fig. 2A. Depicts dose dependent attenuation of TFMPP on Light-Induced Phase Advances. TFMPP (5.0 mg/kg, i.p.) was administered 30 minutes prior to CT 19. At CT 19, hamsters were exposed to 10 minutes of white light (10 minutes of 20 lux).

Results

Phase advances resulting from light stimulation were attenuated by TFMPP in a dose dependent manner. Administration of 5.0 mg/kg TFMPP completely inhibited photic induced phase shifts (-0.04 ± 0.14 hrs) as compared to saline (1.12 ± 0.06 hrs). Figure 2A depicts inhibition of photic induced phase shifts as a function of dose of TFMPP.

Discussion

The data clearly indicate a dose dependent relationship between TFMPP and inhibition of phase shifts. In light of our lack of certainty regarding TFMPP's exact mechanism of action in the SCN, additional data is required in order to elucidate the receptor subtype responsible for these effects. Administration of selective 5-HT antagonists in conjunction with low doses of TFMPP and subsequent assessment of phase shift inhibition should prove interesting.

Experiment 3

Effect of TFMPP on Fos Expression in the Suprachiasmatic Nucleus

Introduction

Light stimulation appears to be a direct cause of c-fos gene transcription, the product of which is Fos protein. The daily pattern of Fos expression in the suprachiasmatic nucleus (SCN) is a function of the external light-dark cycle, and does not seem to be generated endogenously. Interestingly, Fos expression resulting from photic stimulation accompanies phase shifts in circadian activity; likewise, photic induction of Fos protein is also phase dependent. Foe example, Fos levels of animals housed in constant darkness were increased after light stimulation (~300 lux) during their subjective night, but negligible if administered during the subjective day (CT 4-8)[22]. Thus, Fos is a good marker for the occurrence of the mechanism which is responsible for the photic entrainment of the circadian pacemaker.

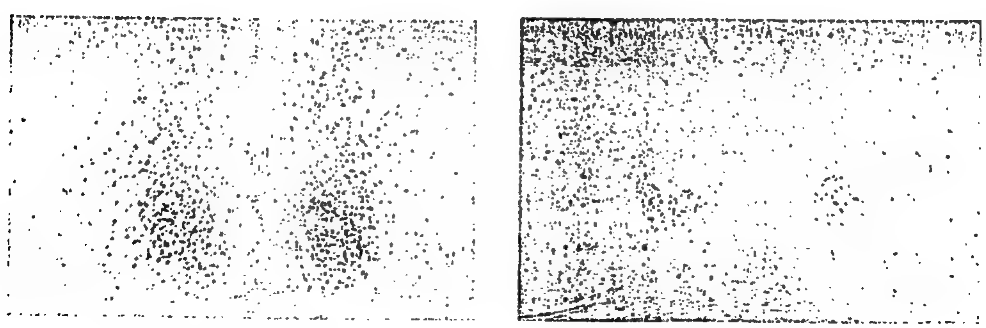
Fos-like immunoreactivity appears selectively in those regions which seem to be involved in photic entrainment of circadian rhythms. Immunoreactive Fos staining is largely restricted to the ventrolateral subdivision of the SCN, a pattern concomitant with the termination of visual inputs[23,24]. No other region receiving retinal input besides the SCN (e.g., the superior colliculus and lateral geniculate nucleus), with the exception of the intergeniculate leaflet, stains positively for Fos. Based in this information, as well as our finding that TFMPP effectively blocked photic induced phase shifts, we would predict a decrease in Fos expression induced by light stimulation in the SCN after the administration of TFMPP.

Methods and Materials

Subjects Adult male Syrian Hamsters Mesocricetus auratus (Charles River, MA) weighing approximately 150g at the time of experiment. Animals were originally housed six to a cage in a vivarium maintained on a 14:10-h light-dark (LD) cycle (lights on at 20:00 hr) for at least 14 days prior to transfer to constant darkness (DD) in individual cages containing 14-cm running wheels. Food and water were supplied ad libitum.

Methods After 10 days in DD, animals received i.p. injections of TFMPP (5.0 mg/kg) or 0.9% saline, 30 min prior to CT 10 under dim red light. At CT 19, animals hen received light stimulation (10 min of 290 lux) and were returned to darkness for 90 min, at which time they were anesthetized and perfused transcardially. Brains were removed and post-fixed in 4% paraformaldehyde at 4°C overnight, then incubated in 0.1M sodium phosphate buffer (pH 7.4) for 24 hours. Coronal sections were cut 70µm thick on a vibratome and then incubated at 4°C overnight in Fos antiserum (c-Fos Ab2@1:2000; Oncogene Science, Manhasset, NY), described previously[25]. Two raters counted immunoreactively stained cells in each SCN; the cell count reported here is an average of the two raters' counts. Stained cells were counted using a reticle on the area corresponding to the SCN.

Light-Induced Fos Expression in the SCN



Vehicle + Light   TFMPP + Light

Fig. 3A. Coronal sections through the SCN region of a hamster injected i.p. with TFMPP (5 mg/kg) 30 min prior to light stimulation (10 min of 20 lux) at CT 19. Sections have been stained for Fos immunoreactivity. Notice the distinctive pattern resulting from TFMPP administration; TFMPP + light resulted in a substantial although anatomically selective decrease in immunoreactive staining for Fos in the SCN as compared to vehicle + light.

Results

Administration of saline injections prior to light stimulation resulted in an abundance of Fos expression in the SCN. Administration of TFMPP resulted in a distinctive pattern of substantially reduced immunoreactive staining for Fos in the SCN (291 ± 426 cells/SCN) as compared to saline (1130 ± 211 cells/SCN) (Fig. 3A).

Discussion

Administration of TFMPP prior to light stimulation resulted in considerable inhibition of Fos-like immunoreactivity in the SCN. In drawing conclusions from these data, it is important to consider the pharmacological specificity of TFMPP. As mentioned earlier (see *Discussion*, Experiment 2), TFMPP binds with some specificity to a number of 5-HT receptor subtypes. Therefore, at this time one cannot confidently attribute TFMPP's inhibition of Fos expression to the 5-HT_{1B} receptor.

Interestingly, the pattern of staining resulting from TFMPP administered prior to light simulation is similar to the pattern of Fos expression commonly seen following administration of MK-801, a NMDA antagonist. Thus, pre-synaptic modulation of the serotonergic system, as well as post-synaptic modulation of the NMDA receptor both result in a distinct pattern of Fos-like immunoreactivity in the SCN. This similarity is notable, as it suggests a common neuronal pathway is responsible for the effects of both serotonergic and NMDA agents; not only regarding the expression of Fos in the SCN, but on light-induced phase shifts as well. Fos expression, as described by many[16,24,25], is a reliable correlate of light-induced phase shifts in circadian activity rhythms. Based on this presumption, the results of this experiment are to some extent supportive of our original hypothesis; that presynaptic serotonergic receptors, although not necessarily of the 1B subtype as originally theorized, seem to modulate excitatory amino acid synaptic transmission to the SCN.

Further investigation is needed to elucidate the neuronal mechanisms underlying this distinct pattern of Fos in the SCN; e.g., identification of the specific pathway through which TFMPP-sensitive receptors project to areas which display Fos like immunoreactivity.

Experiment 4

5-HT_{1B} Receptor Autoradiography in the Suprachiasmatic Region

Introduction

Receptor autoradiography studies have shown a reduction in 5-HT_{1B} binding in the hamster superior colliculus (SC) subsequent to ocular enucleation[13]. Approximately 90% of

retinal afferents project to the contralateral SC[26]; this percentage is closely correlated to the decrease in binding observed following unilateral ocular enucleation. In this study, [^{125}I]iodocyanopindolol, in the presence of isoproterenol, was used to label 5-HT_{1B} receptors in the hamster SCN. Stated again, our hypothesis is that 5-HT_{1B} receptors are located on retinal terminals in the SCN. Thus, destruction of retinal afferents via ocular enucleation should reduce 5-HT_{1B} binding in the SCN.

Groups of animals were subjected to bilateral, unilateral, or sham enucleations. The hamster SCN received approximately 50% of it's retinal innervation from the contralateral eye; therefore, we predict the group receiving unilateral enucleations to show approximately a 50% reduction in binding, and the group receiving bilateral enucleations to show a substantially larger reduction. In addition, we expect to reproduce results similarly obtained with the same experimental manipulations in the superior colliculus[13]. The localized reduction in 5-HT_{1B} receptors in the SCN following destruction of retinal afferents would indicate that 5-HT_{1B} receptors are located on those nerve terminals.

Materials and Methods

Subjects Adult male Syrian hamsters Mesocricetus auratus (Charles River, MA) weighing approximately 100-130 g at the time of purchase. Animals were originally housed 6 to a cage in a vivarium maintained on a 14:10h light-dark cycle (lights on at 20:00 hr) for at least 14 days prior to surgery. Food and water were supplied ad libitum.

Surgical Procedure Hamsters were divided randomly into three groups (n=4) and received either sham (anesthesia only), bilateral, or unilateral enucleations. Subjects were anesthetized with a cocktail of ketaset 125 mg/kg), xylazine (20 mg/kg) and acepromazine maleate 2 mg/kg). Under anesthesia, the optic nerve was cut, a small piece of gelfoam was placed in each orbit and the eyelids sutured shut. Subsequent to recovery from anesthesia, animals were placed into constant darkness in individual cages containing 14-cm running wheels.

Preparation and Development of Autoradiograms Fourteen days following surgery, hamsters were killed by rapid decapitation. Brains were removed and blocked on the coronal plane before being placed on ice and frozen at -80°C. Sections were cut 15mm thick on the coronal plane. Sections were mounted onto subbed slides and stored at -80°C.

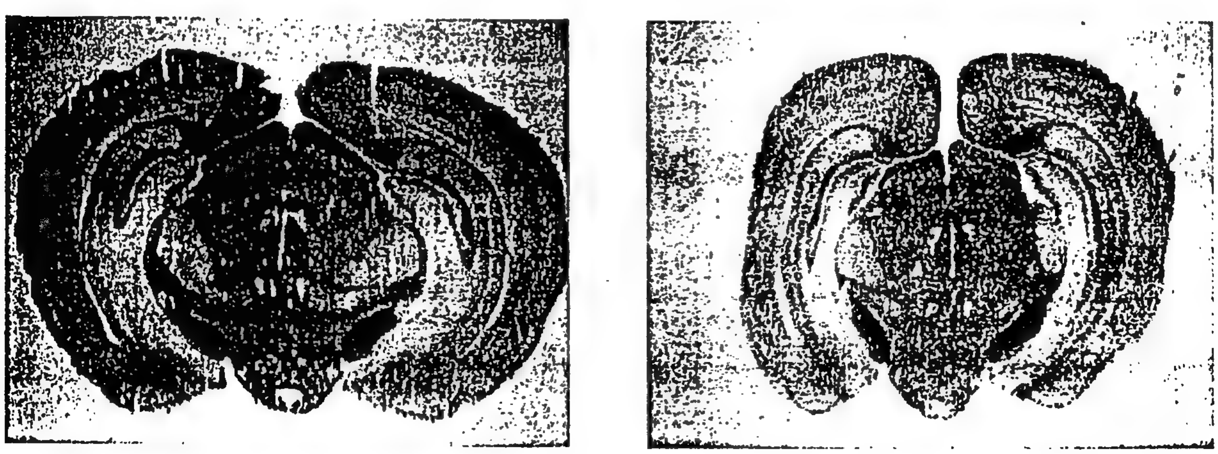
The protocol for the 5-HT_{1B} autoradiographic assay generally followed that of Manaker and Verderame[27]. Frozen sections were allowed to thaw at room temperature and then immersed in ice-cold buffer [50mM tris(hydroxymethyl)aminomethane-HCL, 2.5mM $MgCl_2$, and 20mM isoproterenol, pH 7.4] for 10 min Sections were then incubated in the same buffer containing 100pM [^{125}I]iodocyanopindolol ([^{125}I]ICYP) [specific activity 2200 Ci/mmol] at room temperature for 30 min. Preliminary modifications of the incubation buffer verified that the

addition of 30µg isoproterenol and 20µg 5-HT effectively blocked [^{125}I]ICYP binding to b-adrenergic receptors and all other non-specific binding. Following incubation, sections were washed twice for 5 min in cold buffer, then dipped in ice-cold distilled water in order to remove any remaining buffer. Slides were set on top of a slide warmer and allowed to dry. Finally, sections were apposed to Hyperfilm (Amersham Life Science) for 20-22 hrs. The exposed films were then developed in Kodak D-19 developer for 3-5 min each.

Results

5-HT_{1B} binding in the hamster superior colliculus, as labeled with [^{125}I]ICYP in the presence of isoproterenol, was bilaterally reduced in hamsters receiving bilateral enucleations. In addition, 5-HT_{1B} binding was reduced on the side opposite unilateral enucleations, and no reduction in binding occurred in subjects having received sham treatment. These results replicated previous (see Fig. 4A). Whether or not such reductions took place in the SCN is unclear at this point in time. Figure 4B compares representative SCN sections of a bilateral enucleation and a control labeled with [^{125}I]-ICYP. some reduction in binding in the enucleated animal appears to have taken place in the ventromedial subdivision of the SCN. This reduction appears to be qualitatively different from the control animal; however, a more qualitative assessment is needed for more effective interpretation of results. Systematic differences in binding between unilaterally and bilaterally enucleated animals were not generally observable.

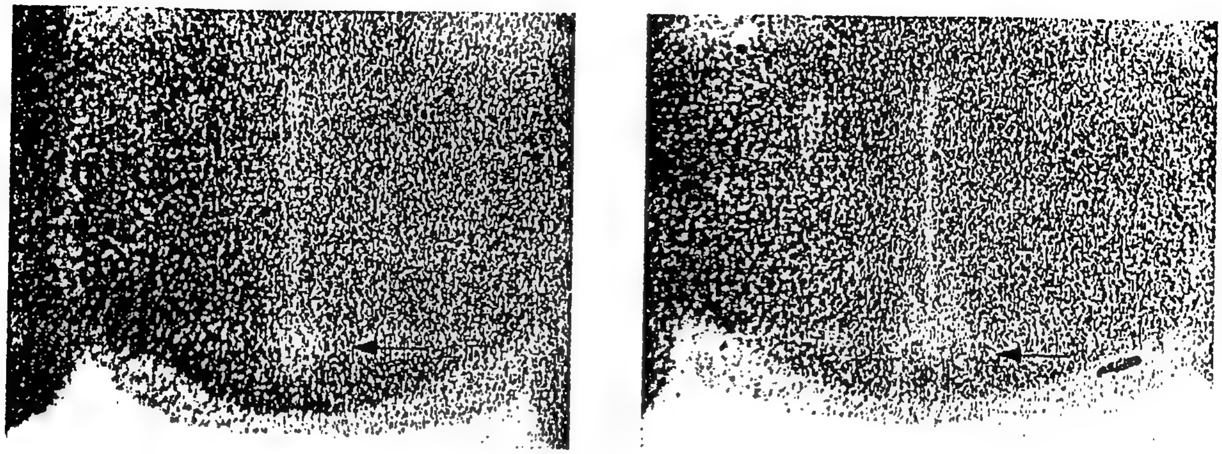
^{125}I-ICYP Binding in the Hamster Superior Colliculus



Sham Enucleation **Unilateral Enucleation**

Fig.4A. Coronal sections of the superior colliculus as labelled for 5-HT1B binding with [125I]ICYP in the presence of isoproterenol. Note decreases in binding in the unilaterally enucleated animal as compared to sham.

^{125}I-ICYP Binding in the Hamster SCN



Sham Enucleation   Bilateral Enucleation

Fig. 4B. Coronal sections of the SCN region labelled for 5-HT1B binding with 125I-ICYP in the presence of isoproterenol. Bilateral ocular enucleations resulted in a small but observable decrease in label throughout the ventromedial region of the SCN as compared to controls.

Discussion

Although no qualitative differences in binding were observed in animals which had undergone unilateral enucleations, results do provide useful information regarding he serotonergic mechanisms involved in the inhibition of light-induced phase advances. The apparent reduction of 5-HT$_{1B}$ binding in the SCN following the destruction of retinal afferents reveals that at least some 5-HT$_{1B}$ receptors are located on axon terminals presynaptic to the SCN. If it is the case that only a small number of 5-HT$_{1B}$ receptors exist on retinal terminals, a 50% reduction in binding as was expected to result from unilateral enucleations might have been an insufficient amount to result in any noticeable decrement; thus bilateral enucleations might have resulted in a more noticeable decrease, as was the case. Moreover, it is possible that the time course between ocular enucleations and tissue removal allowed for receptor upregulation to take place. This rationale is plausible but unlikely, as such upregulation did not appear to take place in the superior colliculus. Finally, it is possible that our hypothesis concerning the location of the 5-HT$_{1B}$ receptor was misguided, in which case it will be necessary to investigate the distribution of other 5-HT receptor subtypes in this region.

References

1.) Moore, R.Y. and Lenn, N.J. (1972) A retinohypothalamic projection in the rat. *J Comp Neurol*, **146**:1-14.

2.) Pickard, G.E. (1980) Morphological characteristics of retinal ganglion cells projexting to the suprachiasmatic nucleus: A horseradish peroxidase study. *Brain Res*, **183**:458-65.

3.) Colwell, C.S. and Menaker, M. (1994) NMDA as well as non-NMDA receptor antagonists can prevent phase shift effects of light on the circadian system of the golden hamster. *J Biol Rhythms*, **7**:125-136.

4.) Cahill, G.M. and Menaker, M. (1988) Effects of excitatory amino acid receptor antagonists on suprachiasmatic nuclei responses to retinohypothalamic volleys. *Brain Res*, **479**:76-82.

5.) Rea, M.A., Michel, A.M. and Lutton, L.M. (1992) Effects of local administration of excitatory amino acids on circadian phase and c-Fos expression in the hamster suprachiasmatic nuclei. *Soc Neurosci Abs*, **18**:3.

6.) Takeuchi, Y., Takashima, M., Katoh, Y., Nishikawa, T. and Takahashi, K. (1991) N-methyl D aspartate and kainate receptors are all involved in the transmission of photic stimulation in the suprachiasmatic nucleus in rats. *Brain Res*, **563**:127-131.

7.) Azmitia, E.C. and Segal, M. (1978) An autoradiographic analysis of the differential ascending projections of the dorsal and median raphe nuclei in the rat. *J Comp Neurol*, **179**:641-668.

8.) Moore, R.Y., Halaris, A.E. and Jones. (1978) Serotonin neurons of the midbrain raphe: ascending projections. *J Comp Neurol*, **180**:417-38.

9.) Miller, J.D. and Fuller, C.A. (1990) The response of the suprachiasmatic neurons of the rat hypothalamus to photic and serotonergic stimulation. *Brain Res*, **515**:155-62.

10. Srkalovic, G., Slim, M. Rea, M.A. and Glass, D.J. (1994) Serotonergic inhibition of extracellular glutamate in the suprachiasmatic nuclear region assessed using in vivo brain microdialysis. *Brain Res*, **656**:302-8.

11.) Boschert, U., Ait. Amara, D., Segu, L and Hen, R. (1994) The mouse 5-hydroxytryptamine$_{1B}$ receptor is localized predominantly on axon terminals. *Neurosci*, 58(1):167-182.

12.) Prosser, R.A., Dean, R.R., Edgar, D.M., Heller, H.L. and Miller, J.D. (1993) Serotonin and the mammalian in vitro circadian system I: In vitro phase shifts by serotonergic agonists and antagonists. *J Biol Rhythms*, **8**:1-16.

13.) Mooney, R.D., Shi, M.Y. and Rhoades, R.W. (1994) Modulation of retinotectal transmission by presynaptic 5-HT$_{1B}$ receptors in the superior cololiculus of the adult hamster. *J Neurophysiol*, **72**:3-13.

14.) Rea, M.A., Buckley,B. and Lutton, L.M. (1993) Local administration of excitatory amino acid antagonists blocks light-induced phase shifts and c-Fos expresson in the hamster SCN. *Am J Physiol*, **34**:1191-98.

15.) Sills, M.A., Wolfe, B.B. and Frazer, A. (1984) Determination of selective and non-selective compounds for the 5-HT$_{1A}$ and 5-HT$_{1B}$ recceptor subtypes in rat frontal cortex. *J Pharmacol Exp Ther*, **247**:1073-81.

16.) Engel, G. Gothert, M., Hoyer, D., Schlicker, E. and Hillenbrand, K. (1986) Identity of inhibitory presynaptic 5-hydroxytryptamine (5-HT) autoreceptors in the rat brain cortex with 5-HT$_{1B}$ binding sites. *Naunyn-Schmiedeberg's Arch Pharmacol*, **332**:1-7.

17.) Conn, P.J. and Sanders-Bush, E. (1987) Relative efficiencies of piperazines at the phosphoinositide hydrolysis-linked serotonergic (5-HT-2 and 5-HT-1c) receptors. *J Pharmacol Exp Ther*, **242**:552-557.

18.) McAllister, G., Charlesworth, A., Snodin, C., Beer, M.S., Noble, A.J., Middlemiss, D.N., Iversen, L.L. and Whiting, P. (1992) Molecular cloning of a serotonin receptor from human brain (5-HT1E): A fifth 5-HT1-like subtype. *Proc Natn Acad Sci*, USA, **89**:5517-21.

19.) Lovernberg, T.W., Erlander, M.G., Baron, B.M., Racker, M. Slone, A.L., Siegel, B.W., Craft, C.M., Burns, J.E., Danielson, P.E. and Sutcliffe, J.G. (1993) Molecular cloning and functional expression of 5-HT1E-like rat and human 5-hydroxytryptamine receptor genes. *Proc Natn Acad Sci,* USA, **90**:2184-2188.

20.) Leonhardt, S., Gorospe, E., Hoffman, B.J. and Teitler, M. (1992) Molecular pharmacological differences in the interaction of serotonin with 5-hydroxytryptamine 1C and 5-hydroxytryptamine 2 receptors. *Molec Pharmacol,* **42**:328-335.

21.) Fouguet, M., Hoyer, D., Pardo, L.A., Parekh, A., Kluxen, F.W., Kalkman, H.O., Stuhmer, W. and Lubbert, H. (1992) Cloning and functional characterization of the rat stomach fundus serotonin receptor. *EMBOJ,* **11**:3481-3487.

22.)Schwartz, W.J., Takeuchi, J., Shannon, W., Davis, E.M. and Aronin, N. (1994) Temporal regulation of light-induced Fos and Fos-like protein expression in the ventrolateral subdivision of the rat SCN. *Neurosci,* **58**(3):573-83.

23.) Rea, M.A. (1992) Different populations of cells in the suprachiasmatic nuclei express c-Fos in association with light-induced phase delay and advances of the free running activity rhythm in hamsters. *Brain Res,* **579**:107-112.

24.) Abe, H., Rusak, B., and Robertson, H.A. (1992) NMDA and non-NMDA receptor antagonists inhibit photic induction of Fos protein in the hamster suprachiasmatic nuclei. *Brain Res Bull,* **28**:831-35.

25.) Rea, M.A. (1989) Light increases Fos-related protein immunoreactivity in the rat suprachiasmatic nuclei. *Brain Res,* **23**:577-581.

26.) Pickard, G.E. and Silverman, A.J. (1981) Direct retinal projections to the hypothalamus, pyriform cortex, and accessory optic nuclei in the golden hamster as demonstrated by a sensitive anterograde horseradish peroxidase technique. *J Comp Neurol,* **211**:65-83.

27.) Manaker, S. and Verderame, H.M. (1990) Organization of serotonin 1A and 1B receptors in the nucleus of the solitary tract. *J Comp Neurol,* **301**:65-83.

NETWORKED HYPERMEDIA FOR DISTANCE LEARNING: TRENDS AND CURRENT USES OF THE WORLD WIDE WEB FOR HYPERMEDIA INSTRUCTION.

Eric O. Riise
Department of Communication Design

California State University, Chico
Chico, CA 95928

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by: Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1995

NETWORKED HYPERMEDIA FOR DISTANCE LEARNING: TRENDS AND CURRENT USES OF THE WORLD WIDE WEB FOR HYPERMEDIA INSTRUCTION.

Eric O. Riise
Department of Communication Design
California State University, Chico

Abstract

The emergence of a global information infrastructure and the use of the World Wide Web for distance learning is gaining popularity. The reasons being an increase in demand for learning and an inexpensive means of delivery. Though current technology is increasing capabilities, availability of the web is not wide spread. Flexibility, interactivity and lack of a fluid authoring system were also areas of weakness.

NETWORKED HYPERMEDIA FOR DISTANCE LEARNING: TRENDS AND CURRENT USES OF THE WORLD WIDE WEB FOR HYPERMEDIA INSTRUCTION.

Eric O. Riise

Introduction

The emergence of a global information infrastructure (U.S. Congress, 1993) and the use of the World Wide Web, also known as "the Web", for distance learning is gaining popularity. The reasons being an increase in demand for learning and an inexpensive means of delivery.

Background

Distance learning is changing for the following reasons: the economical and social contexts have changed as the trend toward lifelong learning increases. As noted in a nation-wide survey recently completed by Washington State University, Pullman, four of five workers polled said they need more education to be successful at work (Technology encourages trend to lifelong learning.. 1995). The need for retraining is also growing as knowledge is rapidly expanding and its lifetime becomes increasingly shorter and shorter. Accordingly, companies need to change, to train and retrain their employed (Peraya, 1995). But how does industry view the Web and its future? At Siggraph '95, one of the largest computer conferences in the world, Silicon Graphics Inc. president and CEO, Tom Jermoluk, called the Web "...the greatest pull for new content opportunities in the history of the computer..." (Web Kudos, 1995). As a means of communication, it is growing "...exponentially and is becoming one of the most important media for global information sharing" (Ibrahim, 1995).

Origins

Hypertext systems are considered to be database applications which provide a unique, non-sequential and flexible method of accessing information through navigation and exploration. The essential features of hypertext are nodes that contain information and links that connect related nodes. It is the linking capability that makes hypertext more powerful than conventional information systems (Balasubramanian, 1994).

Balasubramanian (1994) also indicated that most corporations seem to be interested in making money rather than explore the multimedia aspects and intellectual applications of hypermedia learning systems.

The Web provides a means of instructional delivery that is low in cost and easy to update; it also provides self-directed, self-paced instruction in any topic (Kilby, 1995).

The notion of having the ability to "easily link to other pieces of information, so that only the most important data would be quickly found by a user" has been pursed since the 1940's (Hughes, 1993).

The technology makes such ideas possible is the World Wide Web.

The rise of the Web has been dramatic. According to the most current statistics at the National Science Foundation Network NSFNET The Web has expanded to more than 500,000 current estimated users (GVU's NSFNET Backbone Statistics Page. 1995).

Information available on the Web ranges from links to textual and graphical materials to programs that allow embedding links to self-contained interactive programs within HTML documents.

The Web is a combination of Hypertext Markup Language (HTML) and a transfer protocol for hypertext documents (HTTP). HTTP allows a client application to request data objects from a server.

These objects are identified by a Uniform Resource Locator (URL).

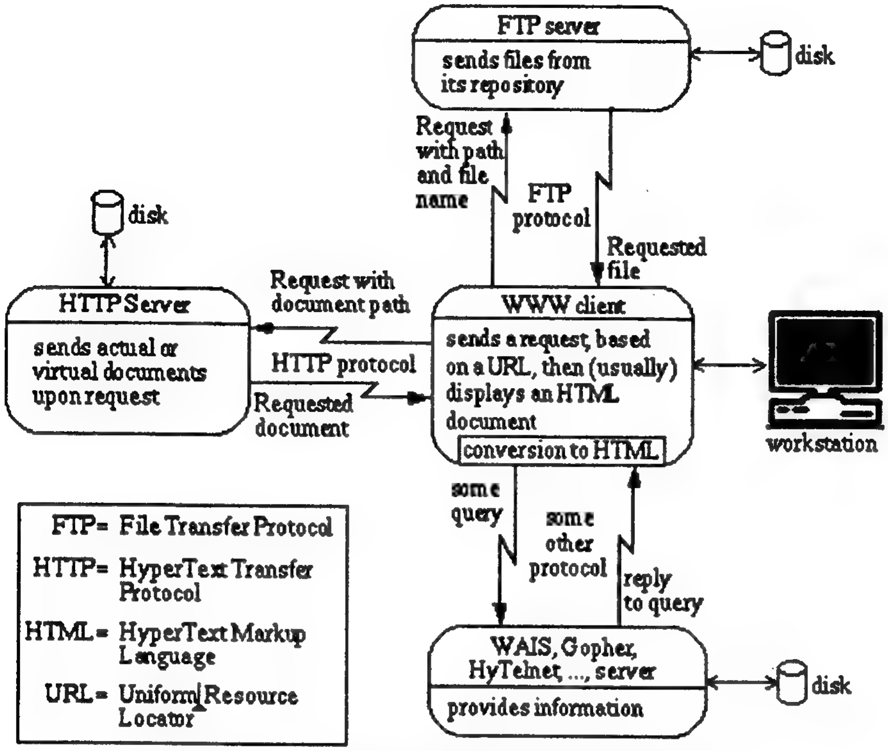


Fig. 1: Client-Server configuration of the World-Wide Web. Depending on the protocol used, the WWW client might need to convert the information to HTML before displaying it.

Figure courtesy Ibrahim, 1995

The concept of hypermedia has been linked to Vannevar Bush and his Memory expander or Memex concept. Bush wrote a paper in 1945 entitled “As We May Think”; in this paper he proposed the Memex. Though not actually implemented, it was “a sort of mechanized private file and library” and “a device in which an individual stores his books, records and communications and which is mechanized so that it may be consulted with exceeding speed and flexibility” (Bush, 1945. p. 101). The information would be stored in memory on an individuals desk. The proposed apparatus also had a scanner to allow the user to input new material as required.

This feature would allow specialists in a field to find information more easily than on paper. "Associative indexing" tied two pieces of information together. The basic idea of which is a provision "whereby any item may be caused at will to select immediately and automatically by another" was thought to be an essential feature of Memex. (Bush, 1945. p. 107).

The next major development was provided by Ted Nelson who originally coined the term "hypertext". Inspired by Bush's grand vision of interlinked data, Nelson is considered the father of the modern version of hypertext. He wrote: "...the future of humanity is at the interactive computer screen, that the new writing and movies will be interactive and interlinked. It will be united by bridges of and we need a world-wide network to deliver it with royalty (Nelson, 1994).

Nelson's vision included a repository of all information with interconnected, or as Nelson puts it "deeply interwingled" links. Though funded by computer software giant Autodesk, his ideas was never realized in a practical sense. Nelson's scheme collapsed in the early 1990s due to its immense projected size, the lack of a clear method of resolving the royalty issue and the inability to procure continued funding.

Web Beginnings

In 1989 applying what he learned from Ted Nelson's ideas Tim Berners-Lee of *Conseil Europeen pour la Recherche Nucleaire,* or CERN, conceived the idea of the World-Wide Web (Relihan, et al, 1994).

The original idea was to create a means for physicists to collaborate on complex physics, and engineering project (CERN-MIT, 1994). By giving researchers a wide-area multimedia information retrieval system providing "universal access to a large universe of documents" (CERN, 1991) collaboration could occur without consideration of space and time.

The next significant development was a graphical-user-interface (GUI) called Mosaic. Developed by a team of students led by Mark Andreeson at the National Center for Supercomputer Applications (NCSA), located at the University of Illinois at Urbana Champaign. Distributed as freeware, this breakthrough browser finally provided a means of presenting multimedia,; text, graphics, sound and video information world-wide in a easy to use, visually pleasing fashion. With Mosaic in place all that was left was the development of content. Among those that have systematically examined the rapidly developing content, available on the Web as it relates to training was in Switzerland.

Uses

The researchers at TECFA ("Technologies de Formation et Apprentissage") a teaching and research unit within the School of Psychology and Education (Faculte de psychologie et des sciences de l'education), University of Geneva navigated though the Web in order to find how it was used Distance Education, and who is using it.

Their findings indicated four basic ways in which it is applied:

1. In a **information desk** mode providing registration and information about various courses at various institutes of higher education.

2. As a **support service** that provides guidance on how to use the Web and Internet, and its underlying information technologies. (e.g. The Internet in the Classroom: a NASA K-12 Internet Initiative).

3. As an **augmentation** tool to support instruction in classrooms. (e.g. Applied Logic Class at Brigham Young University).

4. Finally, there are **online courses** that are delivered at a distance (e.g. Applied Logic Class at Brigham Young University).

The first three categories, though interesting, lie beyond the scope of this paper. The focus will be on some specific examples of the how the Web serves as a delivery means for online courses.

<u>Online Courses</u>

Silicon Graphics CEO Tom Jermoluk stated during his Siggraph '95 speech that "half of our Web business is behind the firewall stuff" The corporate investment is coming on big time (Web Kudos, 1995). Rich Schmeider, Managing Principle Consultant for Education and Training, at Amdahl shared an example of what "behind the firewall stuff" his company is involved in. Schmeider refers to his system as an "intranet" an internal Web integral to providing an enterprise wide solution to their training needs.

With a number of sites scattered across the country, the intranet provides a valuable means for delivering instruction that provides both just-in-time and easily updated instruction.

Another exemplar of the revolutionary delivery systems being developed for the Web is Web Educational Support Tools (WEST) developed at University College Dublin, Ireland. As a comprehensive courseware delivery system, it can be used to distribute course syllabi, assignments, lecture notes, exams, class calendars, multimedia textbooks, to students and trainers worldwide. (WEST, 1995).

This system provides a sophisticated yet easy to use learning environment to support the delivery of training and educational materials to users over a network such as the World Wide Web on the Internet or a corporate local area network (LAN).

It provides a simple messaging system that allows:

- Announcements -messages broadcast to the entire class.
- Message to tutor mode- for student-tutor communications.

As well as:

• Courseware page that provides a class outline and learning objectives.

• Exercise page that allows learners to test their knowledge as they progress through the course.

Among WEST's limitations are lack of easy to use authoring tools, professional looking, androgogically sound effective learning materials.

Mechanisms need to be provided to protect the copyright of course materials.

Incorporation of video conferencing facilities.

Advantages/ Disadvantages

Research has indicated that the advantages of the Web such as :

1. The ability of training centers to distribute knowledge in a timely fashion on a global scale (Kilby, 1995).

2. Reduced mailing costs at training centers. The Web allows distribution of pages without the overhead of printing costs and transport (Peraya, 1995).

3. All information can be corrected and updated immediately for all users from only one server site (Lintz & Tognotti, 1995).

4. Interactivity allows various teaching techniques to be applied incorporating such rich material as text, images, communication between teachers and learners, and between learners (Lintz & Tognotti, 1995).

Other researchers examining the use of the Web for training point out the effectiveness and cost benefits of well-designed CBT training method for multiple trainees. (Kilby, 1995). Web training caries with it the same benefits of CBT with the additional feature of being flexible in the quickly and economically updated when content changes

"The effectiveness and cost benefits in commercial, industrial, and military applications, enhanced by self-paced, any time/anywhere delivery, are superior to any other current training method for multiple trainees." (Kilby, 1995).

CBT applications may not be quickly and economically updated when content changes. Materials for training could easily be out of date by the time CBT was designed, coded, and distributed. Furthermore, producing CBT for multiple playback platforms is always more costly than for a single computer system. Kilby, (1995) pointed out that"... disadvantages may be overcome by designing CBT for easy modification and universal playback and by choosing production and distribution methods that minimize delivery delays". As a solution to these problems, the Web is an ideal thus web-based training will inevitably transform CBT into a timely, universally accessible training system." (Kilby, 1995).

Issues to be addressed

User Interface Issues include ease of navigation and a systematic approach to usability testing and evaluation of hypertext systems. Balasubramanian (1994) has suggested that navigation through a hypertext database must be at least as rich as those available in books.

Evaluation of hypertext systems must take into account the targeted task domain, the typical user population, and the desired outcomes of navigation. "The World Wide Web-as it currently exists-cannot possibly succeed on the scale now being widely predicted. Web browsers, Web servers, home pages and HTML-coding courses will all, inevitably, hit the wall. It is extremely unlikely that any substantial free-standing business will ever be built based on today's Web.

Evolving trends

The evolving trends in the use of the Web for training include the incorporation of real time motion video , three dimensional representation, and the implementation of small downloadable programs.

Motion video via the Web is being constantly improved. A good example is StreamWorks. StreamWorks is based on the concept of "streaming media" without having to download the audio or video file to a local disk drive first. Users also can

view or hear the information as it is being transmitted, rather than waiting for a file transfer.

Virtual Reality Modeling Language (VRML) provides a means of describing three dimensional space so the user can maneuver in a near 3D environment (Kohlhepp, 1995). Using this environment for modeling complex objects a learner might study the circulatory system from a blood cells perspective or take advantage of a growing body of 3D scientific visualization tools and images.

A further evolving enhancement to the Web is the addition of real-time animation and interaction through in-line applications also known as "applets". Using applets in Web training materials would allow the learner to progress as mastery is gained in progressively more difficult modules.

Conclusion

With the evolution of the Web as a tool for distance learning and training on a global scale coupled with continually evolving possibilities, future uses will continue to expand. The need for more comprehensive research is an overstatement. More rigorous studies in the appropriate use of the Web as an instructional media delivery system is required as well as a robust evaluation tool to assess the weakness and strengths of the World Wide Web for training in all sectors-military and civilian.

Bibliography

Balasubramanian, V. (1994). The State of the Art Review on Hypermedia Issues And Applications. [Online] http://www.isg.sfu.ca/~duchier/misc/hypertext_review/chapter9.html

Baron, T. (1995, August 21). Xing Offers Video, Audio Over 'Net Communications Week. 571, 14.

Bush, V. (July 1945). As We May Think. The Atlantic Monthly. Also [Online] http://www.isg.sfu.ca/~duchier/misc/vbush/

CERN. (1991). The Policy of the WWW Project at CERN. [Online] http://www.w3.org/hypertext/WWW/Policy.html

CERN-MIT. (1994). Joint CERN-MIT Initiative. [Online] http://www.w3.org/hypertext/WWW/Consortium/Prospectus/Hosts.html

Web Kudos. (1995, August 14). Interactive Age. 221, 50.

Hughes, K. (1994). Entering the World-Wide Web: A Guide to Cyberspace. [Online] http://www.eit.com/web/www.guide/

GVU's NSFNET Backbone Statistics Page. (1995). [Online] http://www.cc.gatech.edu/gvu/stats/NSF/merit.html

Ibrahim B. (1995). Distance learning with the World-Wide Web. [Online] http://cuiwww.unige.ch/eao/www/OpenDistLearning/paper.html

Kilby, T. (1995). Web Based Training. [Online] http://www.clark.net/pub/nractive/training.html

Kohlhepp, R. (1995, August 1). Next generation Web browsing, Network Computing , 609, 48.

Lintz, M & Tognotti, S. (1995. Distance Education on the WWW [Online] http://tecfa.unige.ch/~lintz/staf14/staf14+ST/projet.html#part1

Peraya, D. (1995). Distance Education and the WWW [Online] http://tecfa.unige.ch:80/edu-ws94/contrib/peraya.fm.html

Portway, P., & Lane, C. (1994). Teleconferencing and Distance Learning (2nd ed.). San Ramon, CA: Applied Business teleCommunications.

Relihan, L., Cahill T. & Hinchey M. (1994). Untangling the World-Wide Web. [Online] http://itdsrv1.ul.ie/Research/WWW/utwww

Stahlman, M. (1995, September 11). What is wrong with Web today?, Computer Reseller News, 648, 18.

Technology encourages trend to lifelong learning, study says. (1995, September 17). Seattle Times-Intelligencer, p. C12.

U.S. Congress, O. T. A. (1993). Advanced Network Technology-Background Paper. Washington D.C.: USGPO.

Web Educational Support Tools Ltd. (WEST), (1995). An Overview of WEST [Online] http://west.ucd.ie/demo/overview.html

Wilson, A. (1992). Interact '92: Integrated Services Digital Network in Education. Project, California State University, Chico.

Zeltser L. (1995). World-Wide Web: Origins and Beyond. [Online] http://homepage.seas.upenn.edu/~lzeltser/WWW/index.html

SOLID PHASE MICROEXTRACTION FOR MONITORING WATER SOLUBLE JET FUEL COMPONENTS IN GROUND WATER

Jason P. Ritter
Graduate Student
Clarkson University

Clarkson University
Department of Chemistry
Potsdam, NY 13699

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force of Scientific Research
Boiling Air Force Base Washington DC

and

Armstrong Laboratory

September 1995

SOLID PHASE MICROEXTRACTION FOR MONITORING WATER SOLUBLE JET FUEL COMPONENTS IN GROUND WATER

Jason P. Ritter
Graduate Student
Department of Chemistry
Clarkson University

Abstract: The water soluble fractions of aviation jet fuels were examined using solid phase micro extraction. Gas chromatographic profiles of microextracts obtained from the headspace of water samples contaminated by neat jet fuels revealed that each jet fuel water solvable fraction possessed a characteristic profile. This suggests that the fuel contaminant can be identified from the gas chromatograms of the dissolved hydrocarbons.

SOLID PHASE MICROEXTRACTION FOR MONITORING WATER SOLUBLE JET FUEL COMPONENTS IN GROUND WATER

Jason P. Ritter

Introduction

A significant source of hydrocarbon contamination in ground water is due to jet fuel spill and leaks. The spills and leaks that occur on military property include motor gasoline, diesel fuel and aviation turbine fuels. Aviation turbine fuels are used by the United States Air Force, Army, Navy, and Marines. These fuels are JP-4, formally used by the Air Force and Army for flight operations within the continental United States; JP-8, which has replaced JP-4 as the standard turbine aviation fuel; JPTS and JP-7, used by the Air Force for special high altitude flights; JP-5, used by the Navy aboard ships; and aviation gasoline which is used for aircraft with piston engines. Jet-A is another type of jet fuel which is used by civilian air crafts.

Puyer, et. al (1) demonstrated that the water soluble fraction of JP-4 fuel consists primarily of alkyl derivatives of benzene and naphthalene. These analyses were performed by equilibrating pure water with the fuel, passing the aqueous phase through solid phase extraction (SPE) cartridges loaded with C-18 modified silica , and then collecting the extracted organics with an organic solvent to yield a sample amendable to analysis by gas chromatography (GC). Mayfield and Henley (2) employed a modification of Puyear's technique to generate water soluble profiles of various jet fuels and demonstrated that an identification of the various fuel classes could be done. This identification would employ pattern recognition techniques to the profiles of the water soluble fractions of the jet fuels.

Aruther and Pawliszyn (3) developed a microscale analytical technique for sampling samples out of ground water called solid phase micro extraction (SPME). The SPME device is designed to be used with gas chromatography. An SPME device consists of a small rod or fiber of an extractive material housed in a septum penetrating needle and attached to a plunger or other device which can extend the rod for extraction or desorbtion. The device is usually built from a gas chromatograph syringe of the plunge in needle style with the extractive rod attached to the syringe plunger. Typically, the extractive rod consists of a piece of fused silica optical fiber with a thin coating of organic polymer, such as polydimethylsiloxane. The sample is extracted by introducing the device and extending the extractive rod into the sample. The SPME device is left in the sample long enough for the analytes to reach an equilibrium between the headspace (of the sample), aqueous phase (of

the sample) and the organic fiber. Following the extraction period, the rod is withdrawn from the sample into the housing of the needle. The needle is then inserted into the injection port of a gas chromatograph, so that all of the analytes dissolved in the fiber will leave the fiber to enter the heated carrier gas (4).

SPME is very sensitive means of analysis, this is due to the fact that all of sample extract is used for analysis. Other forms of extraction such as SPE only use part of the sample extract for analysis. As defined in EPA method 624 the SPME technique has been successfully applied to the analysis of volatile compounds. This technique has also been able to follow EPA method 625 for extractable compounds. Typical sensitivities obtained are competitive with those of purge and trap analysis for volatile compounds. SPME has a greater analysis sensitivity than the approved EPA method 625 for liquid-liquid extractions (5). SPME has been used to analyze surface and ground water samples for benzene derivatives (6,7). Volatile and semivolatile analytes in food and beverages can also be analyzed by SPME/GC (8).

Although SPME has found much application in direct extraction from aqueous samples, the technique can also be applied to headspace and other vapor phase samples (6). Headspace SPME extractions have been reported for food and beverage components. The headspace technique can be advantageous when either the volatility of the analytes permits a headspace determination or where undesirable components in the bulk sample make direct insertion into the aqueous phase with the SPME device undesirable.

Two types of detectors can be used with the SPME device in a GC, are a flame ionization detector (FID) or a mass spectrometer (MS). In the present work both the SPME/GC-FID and SPME/GC-MS were used to analyze the water soluble fraction of the jet fuel. The SPME technique is expected to provide a rapid extraction technique for these compounds with adequate response over the entire volatility range of substances found in the water soluble fractions of jet fuels.

Experimental

SPME procedures were carried out using commercial SPME holders for manual operation (Supelco). Fused silica extraction fibers with 100 micron polydimethylsiloxane coatings were used for the extractions. Analyses were carried out using either an HP-5890 gas chromatograph/flame ionization detector (GC/FID) or a gas chromatograph/mass spectrometer (GC/MS) consisting of an HP-5890 GC combined with an HP-5970 mass selective detector. Both the GC/FID and GC/MS were equipped with a split/splitless injection ports. For all of the injections, a split injection was used because it was sensitive enough to show all the compounds that a splitless injection did and gave a

enough to show all the compounds that a splitless injection did and gave abetter resolution than the splitless injection.

Calibration standards were generated by dissolving selected analytes at known concentrations in methanol, and then spiking a pure water aliquot with a portion of the methanoic solution. The volume of the spike was kept low, in order to maximize partitioning of the target analyte from liquid to the SPME polymeric coating. Various aromatic hydrocarbons were used as standards. For quantitative experiments d_{10}-ethylbenzene was used as an internal standard. Headspace and direct insertion SPME were compared by analysis of calibration solutions using a GC/FID. The analyses were performed with a 30 meter fused silica capillary column, .025mm in diameter, and coated with 1 micron of a bonded and cross linked nonpolar stationary phase which consisted of 5% phenyl-substituted polymethylsiloxane (DB-5, J&W Scientific). The GC oven was initially set at 40°C for 3 minutes, the temperature was ramped at 10°C/minute to a final temperature of 200°C and a hold time of 0 minutes. Helium was the carrier gas used and was held at a constant pressure of 15 psi. Samples of water saturated with jet fuel components were generated by equilibrating neat jet fuels with water in special equilibration flasks developed for hydrocarbon solubility determinations by Burris and MacIntyre (9). Aliquots of the fuel saturated water were removed from the equilibration flasks by entering air into the system (the air entering the system displaced the water that was in the system) These aliquots were placed in 40 ml VOA vials with one-hole screw caps and teflon-faced septa. Prior to the introduction of the sample, a micro-stirring bar was placed in the vial to permit the sample to be stirred during the SPME sampling procedure. After the samples had been removed from the equilibration flasks and placed in VOA vials, the vials were capped. Methanoic solutions of d_{10}-ethylbenzene which served as an internal standard, were injected through the vial septa.

Fuel-saturated water samples from the laboratory equilibration experiments were analyzed by GC/FID. However, the GC oven program differed from that used to initially compare headspace and direct insertion SPME. The oven temperature was initially -10°C for 3 minutes, ramped at 10°C/minute to 250°C, with a final holding time of 6 minutes.

Samples requiring peak identification were run on a GC/MS system. This instrument was equipped with an identical fused silica capillary column and split/splitless injection ports used in the GC experiments. Detection was provided by an HP-5970 MS, consisting of a conential quadropole mass spectrometer with an electron impact ion source. The MS was coupled to the GC by a direct inlet interface. The GC was temperature programmed from -10°C for 3 minutes, ramped at 10°C/minute to 250°C, with a final holding time of 6 minutes. The mass spectrometer was scanned from 35 to 500 amu.

Results and discussion

Figure 1 shows a plot of peak area as a function of sampling time for two compounds: benzene (Figure 1a) and naphthalene (Figure 1b). The liquid sample used in this experiment was prepared by dissolving enough benzene and naphthalene into methanol so that when the methanol was added to 30 ml of deionized water in a VOA vial both of these substances were at a concentration of 800 ppb. The headspace of the solution was sampled with the fiber. The dissolved aromatics were also sampled in the aqueous phase. This was done by submerging the entire length of the fiber into the water. The fiber was exposed to the sample for fifteen minutes. The analytes gave a representative response of there concentrations in the sample when a fifteen minute exposure time was used. Because the response was higher at shorter sampling times (see Figure 1b), headspace sampling appears to be preferable over direct insertion for higher molecular weight aromatics. Figure 2 shows a total ion chromatogram of a water soluble fraction of jet fuels. The compounds that are in the water sample are listed in the caption for figure 2. The headspace SPME apparently gives a comparable gas chromatogram to that of direct aqueous insertion of the fiber SPME. Headspace SPME and direct aqueous insertion SPME gave equivalent results for a water sample that was saturated with the water soluble components from a JP-8 jet fuel (see Figure 3). Both sampling regimes are expected to give excellent results for jet fuels.

An experiment was also performed to determine the differences between headspace and direct aqueous insertion. This experiment was also used to determine which method was more reproducible. A solution containing 400 ppb of benzene, toluene, o-xylene, p-xylene, and naphthalene was prepared in a VOA vial. Six sample extraction were collected via SPME headspace and SPME direct aqueous insertion. All of the sample extractions were analyzed with a GC/FID, and raw peak areas were generated for all of the compounds. Reproducibility of the sampling extractions was computed by dividing the mean by the standard error mean. The results of this study suggest that the two sampling regimes possess comparable precision. Apparently, small variations in the stirring rate do not significantly affect the rate of mass transfer of the solute to the fiber .

The transport of sample analyte from the aqueous solution to the directly inserted fiber is controlled by either analyte absorption and by the stirring rate. Transport of the analyte in the headspace is controlled by both liquid phase mass transport to the liquid's surface and by gas phase diffusion within the headspace to the sampling fiber. Since liquid phase transport is slower than gas phase diffusion, liquid phase transport may be the rate controlling in both regimes. Finally, the SPME is an integrating device, accumulating analyte during the entire sampling period, so the

device, accumulating analyte during the entire sampling period, so the variance of sampling can be probably reduce by more rapid transport.

Figure 4 shows calibration curves for benzene, toluene, ethylbenzene, and the xylenes (the BTEX compounds) developed, using both headspace SPME and direct aqueous insertion SPME. The concentrations used for each compound were 50, 100, 400, 800, and 1000 ppb. The internal standard used was $D-_{10}$ethylbenzene and was at a constant concentration of 100 ppb. The data was plotted as a response ratio (target analyte divided by the internal standard peak) versus concentration ratio (concentration of target analyte divided by the concentration of the internal standard). A previous study (10) performed at Armstrong labs at Tyndall Air Force Base showed that the an internal standard is necessary to ensure a reliable calibration with SPME (10), since absolute peak areas of identical samples can vary significantly when SPME is used as a sampling device. The higher detection limits observed for benzene and toluene from the calibration curve are probably due to the lower solubility of these compounds in the polymeric coating material. It is evident that either headspace or direct aqueous insertion SPME can be used to determine the concentration of dissolved jet fuel hydrocarbons in water. Although both types of sampling perform about the same, it was decided to use headspace SPME as the method for analysis of jet fuels in ground water, because of the ease and convenience of using this procedure for recovered spill samples.

Figure 5 shows GC profiles representative of water samples contaminated by neat jet fuel samples. One hundred and eighty simulated water samples were prepared from one hundred jet fuel samples, and the GC/FID profiles shown are representative of each jet fuel class. The main goal of this experiment was to see if jet fuel classes could be determined by the hydrocarbons present. It is apparent from the figure 5 that each jet fuel has a characteristic profile. This suggests that fuel types can be determined from there water soluble fraction in contaminated ground water via SPME headspace coupled with GC/FID. The characteristic nature of the GC/FID data also suggests application of pattern recognition methods to identify fingerprint patterns in the data that is characteristic of fuel type (11).

Figures 6 and 7 show a GC/FID of simulated fuel samples prepared from some of the very same fuel samples used in the SPME sampling experiment described in figure 5. The gas chromatograms of these simulated water samples analyzed by SPE were obtained from a 1991 study (2). In this study, an identical 30 meter capillary column was used to perform the GC runs, but the temperature programming rate of the HP-5890 gas chromatograph was 5°C per minute instead of 10°C per minute which was used in this experiment. A slower temperature programming rate was used in the 1991 study, in order to ensure a

higher resolution. When examining figures 6 and 7, it is evident that more compounds (i.e. peaks) are present in the SPME profiles than in corresponding profiles. This suggests that SPME is a better method for extracting jet fuel components from ground water than SPE. It is very likely that SPME is superior to other conventional methods of extraction for sampling fuel materials recovered from ground water. SPME is also more convenient than SPE techniques for sampling organics from aqueous samples.

As noted earlier the GC was run in the split mode. Due to the fact that a splitting ratio of thirty to one was used, one can conclude that the combination of headspace SPME and gas chromatography constitutes a powerful method for detecting fuel materials at low levels in the environment.

Conclusion

GC/FID profiles of the water soluble fraction of aviation jet fuel were obtained using headspace SPME/GC. There are four advantages using SPME for monitoring jet fuels (1) GC profiles characteristic of fuel type can be obtained so fuel spill identification is possible, (2) quantification of the water soluble components of jet fuels is also possible with SPME, (3) headspace SPME is able to provide a more representative sampling of the organics present in the water soluble fraction of jet fuels than conventional extraction techniques such as SPE, and (4) no solvents are needed with SPME so solvent use and waste can be cut down.

Acknowledgements

JR acknowledges the financial support of the Air Force Office of Scientific Research through their Summer Graduate Student Research Program. References to specific trade names or products does not constitute an endorsement on behalf of the United States Air Force.

Table 1:Reproducibility

Compound	Headspace	Direact Aqueous Insertion
Benzene	27.9	81.7
Toluene	44.4	85.4
o-Xylene	26.8	24.6
p-Xylene	47.2	30.9
Naphthalene	121.9	19.5

Figure Captions

Figure 1. GC peak area versus sampling time for the SPME fiber. Each data point is an average of triplicate determinations. Successive points for a particular response curve are connected by a straight line.

Figure 2. Total ion chromatogram of a water sample prepared from compounds representative of the water solvable fraction of an aviation jet fuel. 2a) Headspace SPME, and 2b) direct insertion SPME. 1) Benzene, 2) Toluene, 3) Ethylbenzene, 4) p-Xylene, 5) Propylbenzene, 6) 1,3,5-trimethylbenzene, 7) Siloxane (from SPME fiber) 8) 1,2,4-trimethylbenzene, 9) 1,2,3 trimethylbenzene, 10) diethylbenzene, 11) diethylbenzene, 12) 1,2,3,5 tetramethyl-benzene, 13) Siloxane (from SPME fiber), 14) Naphthalene, 15) Siloxane (from SPME fiber), 16) 1-methylnaphthalene, and 17) 1-ethylnaphthalene. Peak identification was made by the retention time and by matching the observed fragmentation patterns in a mass spectral library.

Figure 3. Total ion chromatogram of a JP-8 fuel. a) Headspace SPME and b) direct aqueous insertion SPME.

Figure 4. Calibration curves for the BTEX compounds obtained by headspace and direct aqueous insertion SPME. For each analyte, adjacent data points are connected by a straight line.

Figure 5. GC profiles of water samples contaminated by jet fuels: JP-4, Jet-A, JP-7, JPTS, JP-5, AND JP-8. Headspace SPME was used to obtain a representative sampling of the dissolved hydrocarbons.

Figure 6. GC profiles of water samples contaminated by jet fuels: JP-4 and Jet-A. Both the SPME and SPE techniques were used to obtain a representative sampling of the dissolved hydrocarbons. The longer retention times in the SPE experiment are attributed to a slower temperature programming rate.

Figure 7. GC profiles of water samples contaminated by jet fuels: JP-5 and JPTS. Both the SPME and SPE techniques were used to obtain a representative sampling of the dissolved hydrocarbons. The longer retention times in the SPE experiment are attributed to a slower temperature programming rate.

Figure 1A

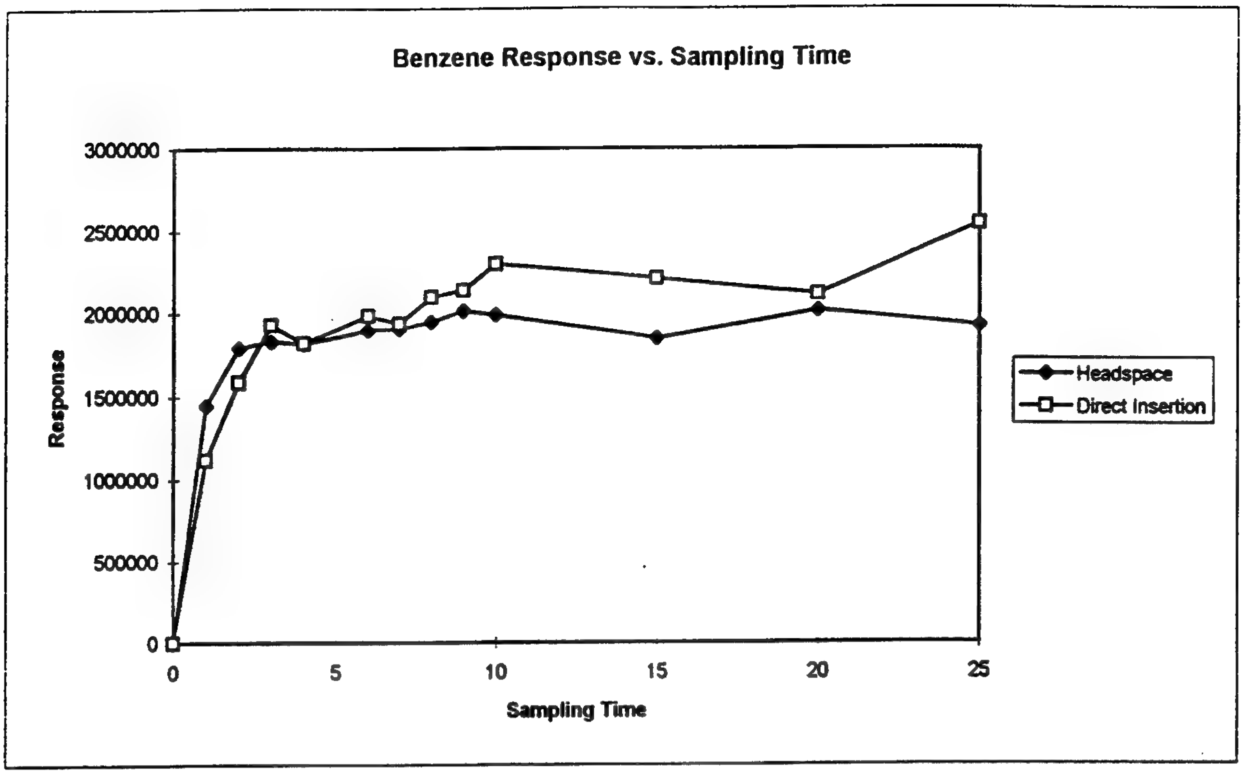


Figure 1B

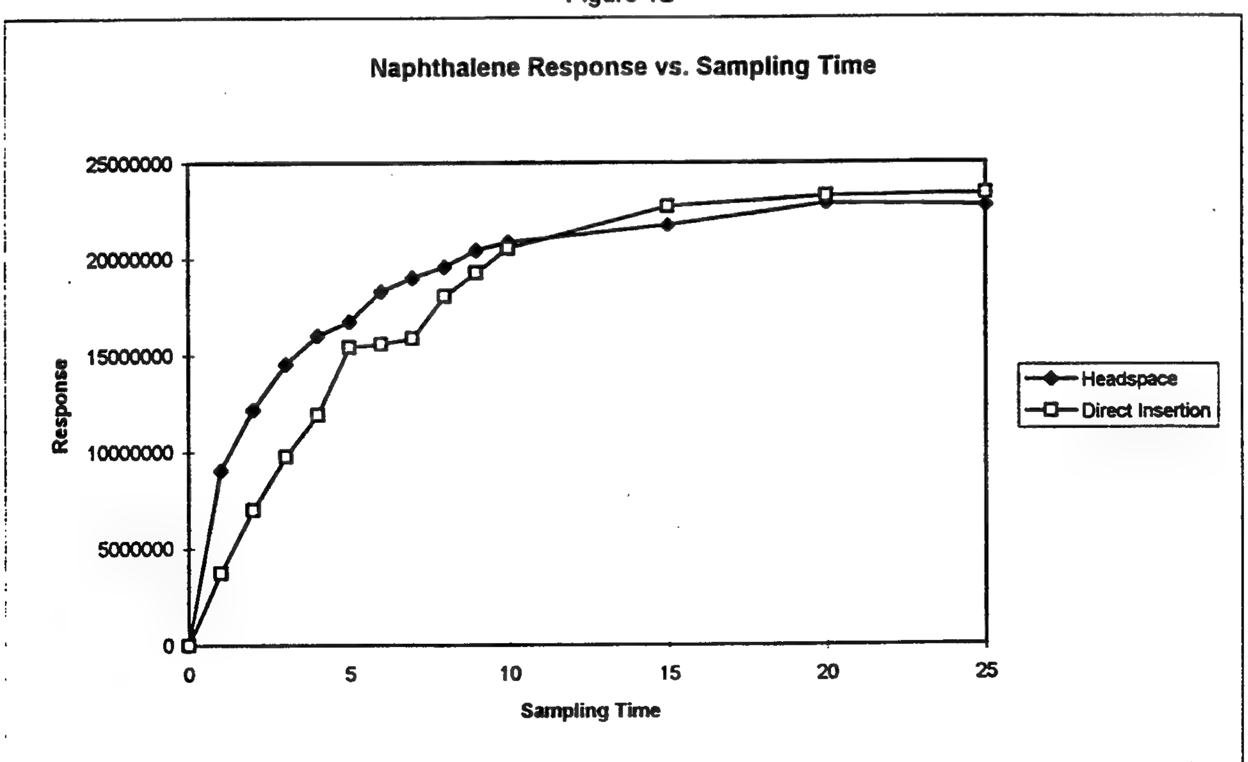


Figure 2

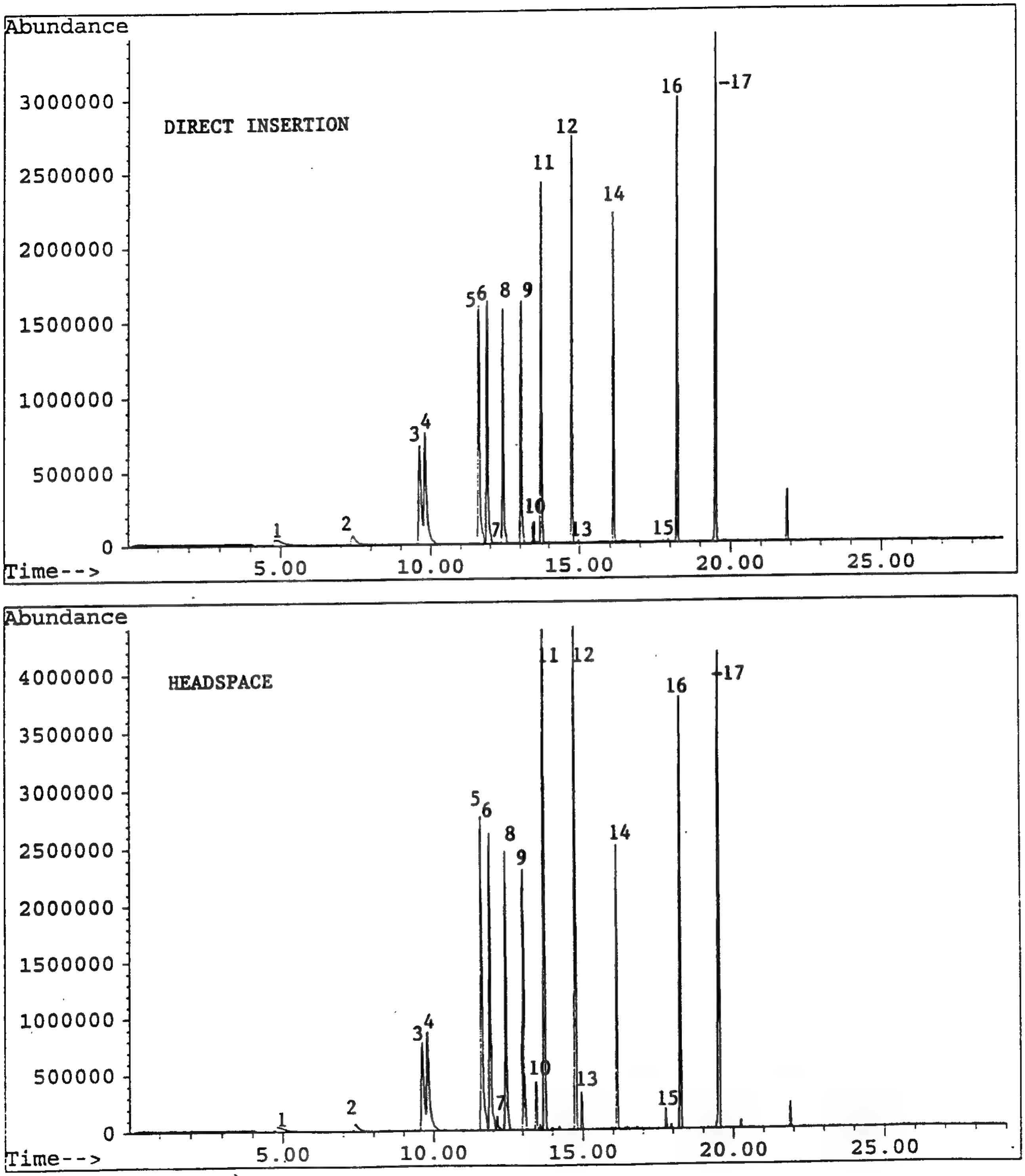

Abundance
DIRECT INSERTION
3000000
2500000
2000000
1500000
1000000
500000
0
Time-->
5.00
10.00
15.00
20.00
25.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
-17
Abundance
HEADSPACE
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
Time-->
5.00
10.00
15.00
20.00
25.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

Figure 3

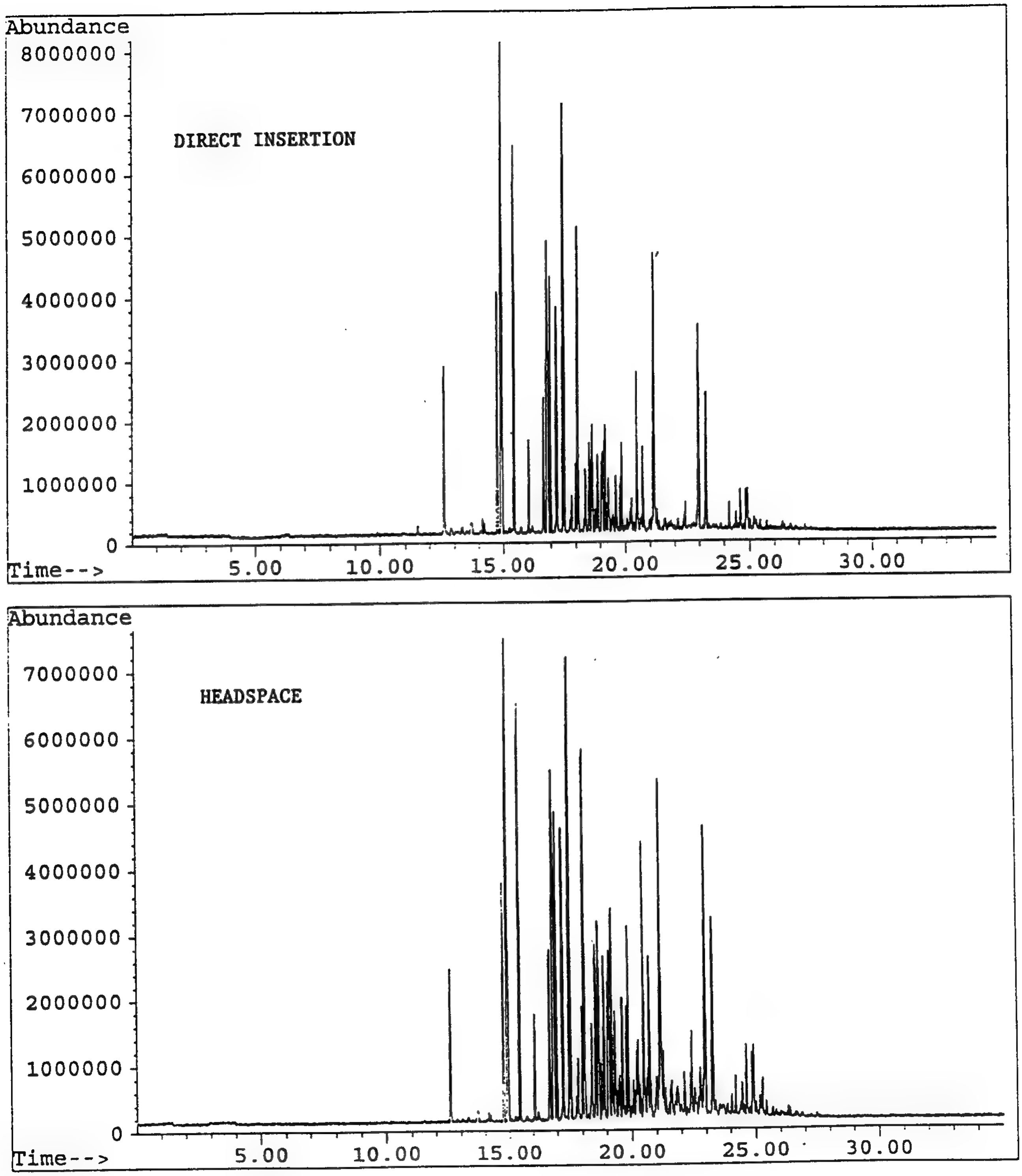

Abundance
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
DIRECT INSERTION
Time-->
5.00
10.00
15.00
20.00
25.00
30.00
Abundance
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
HEADSPACE
Time-->
5.00
10.00
15.00
20.00
25.00
30.00

FIGURE 4A

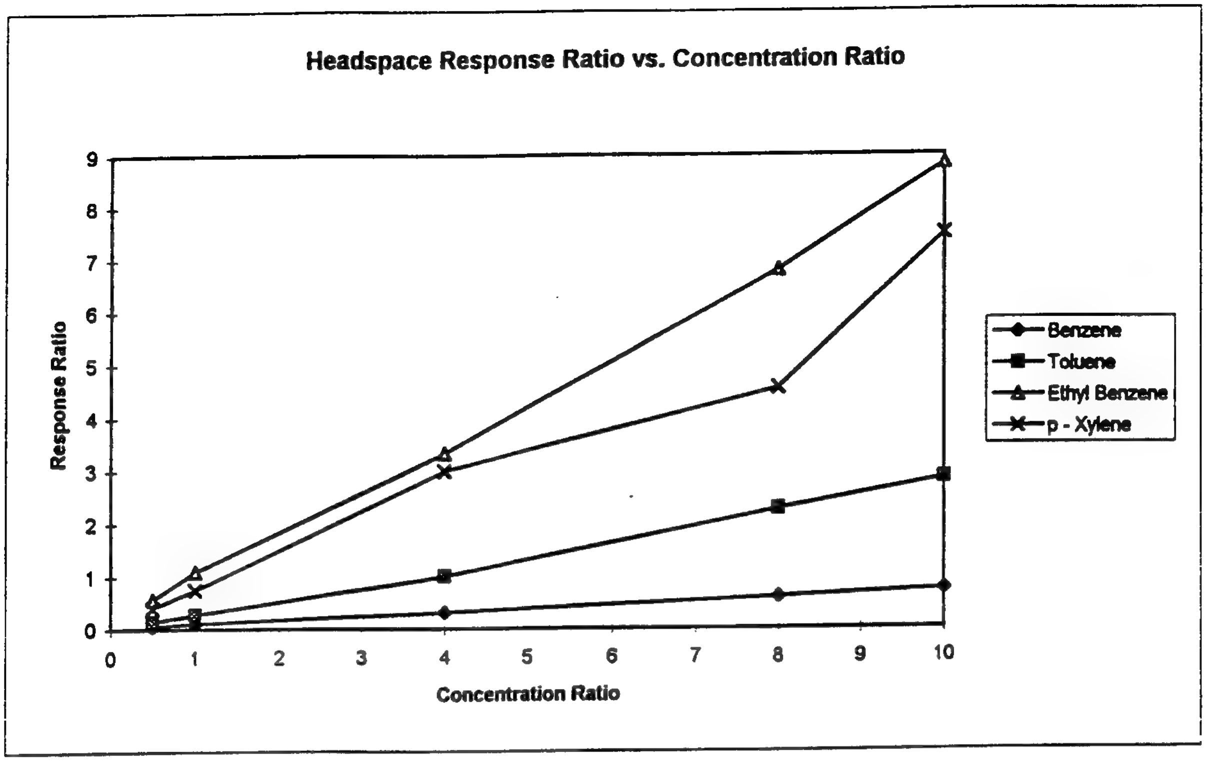


FIGURE 4B

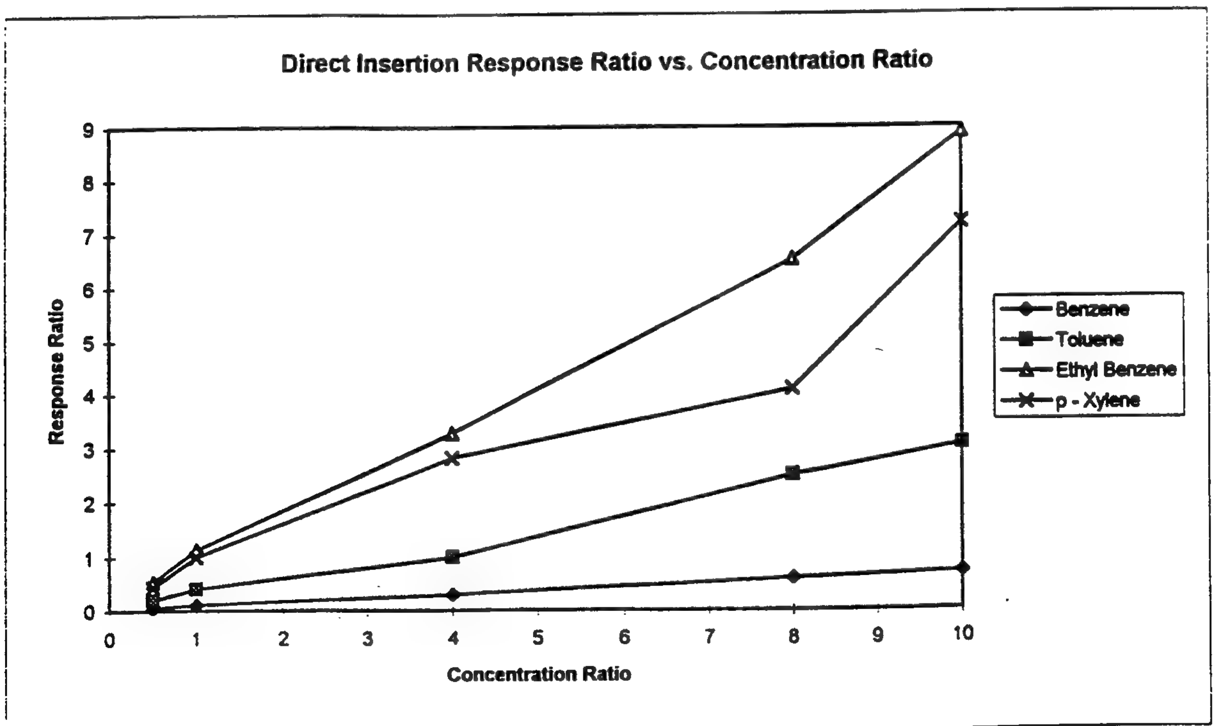


Figure 5

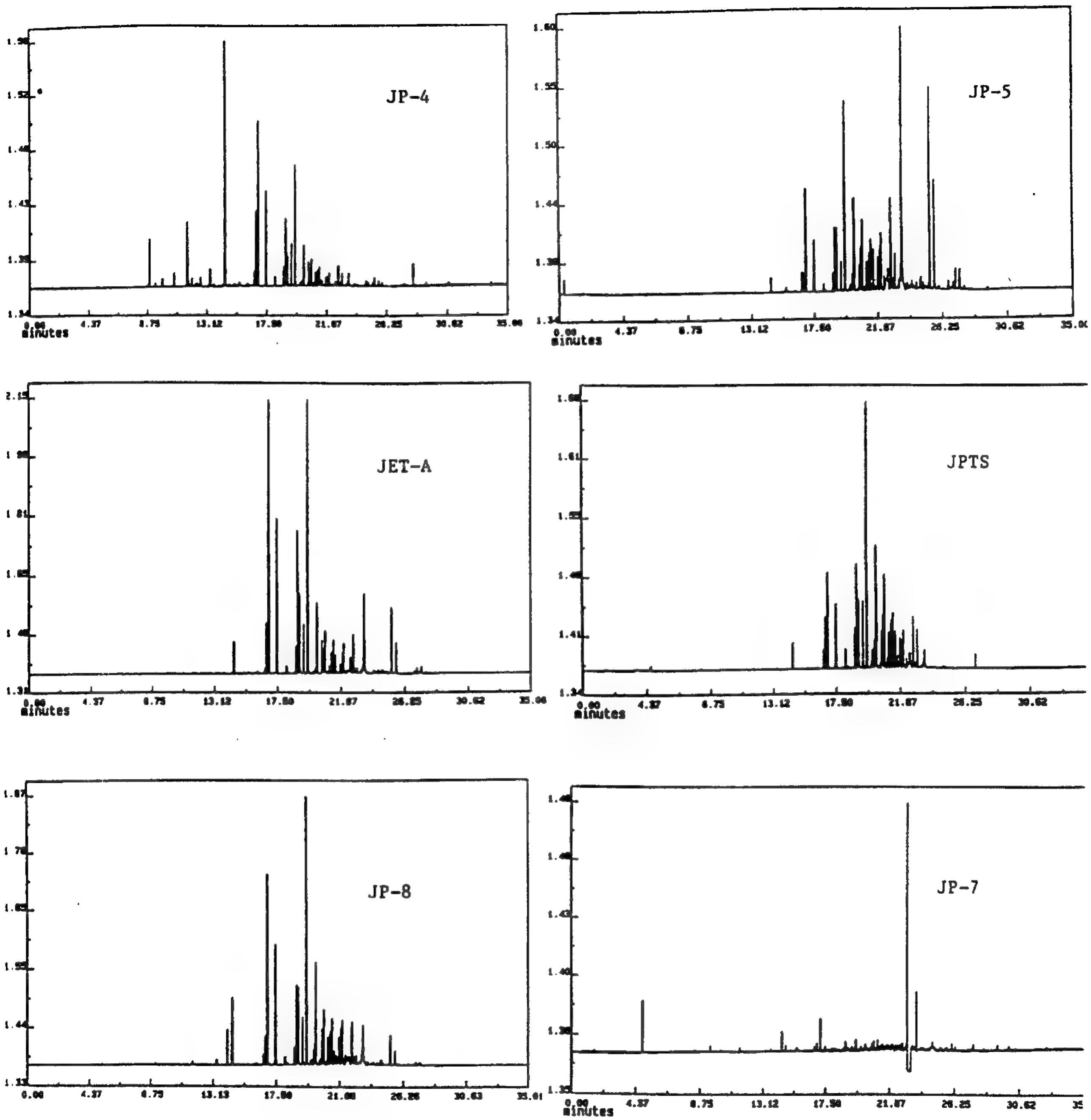


Figure 6

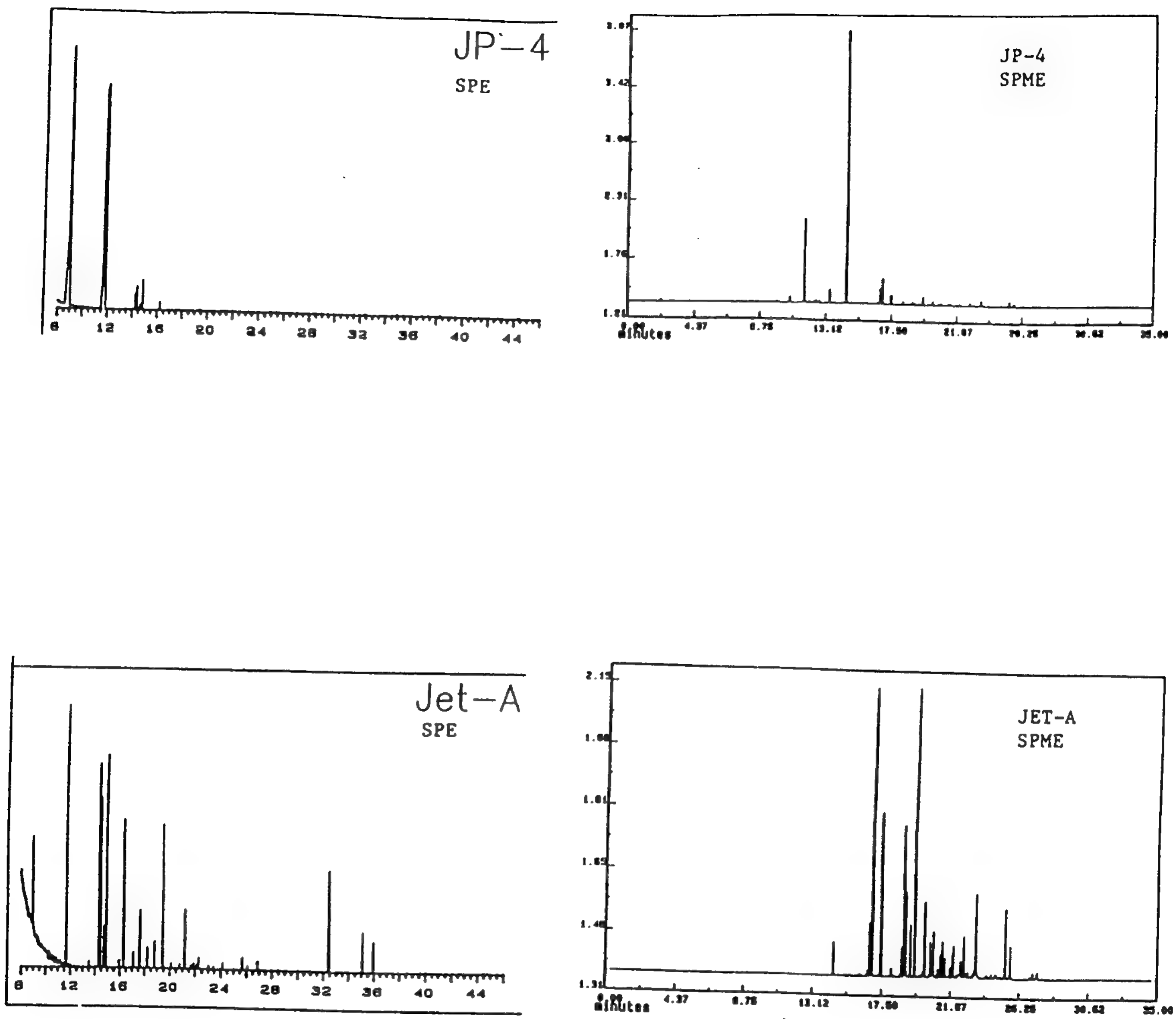
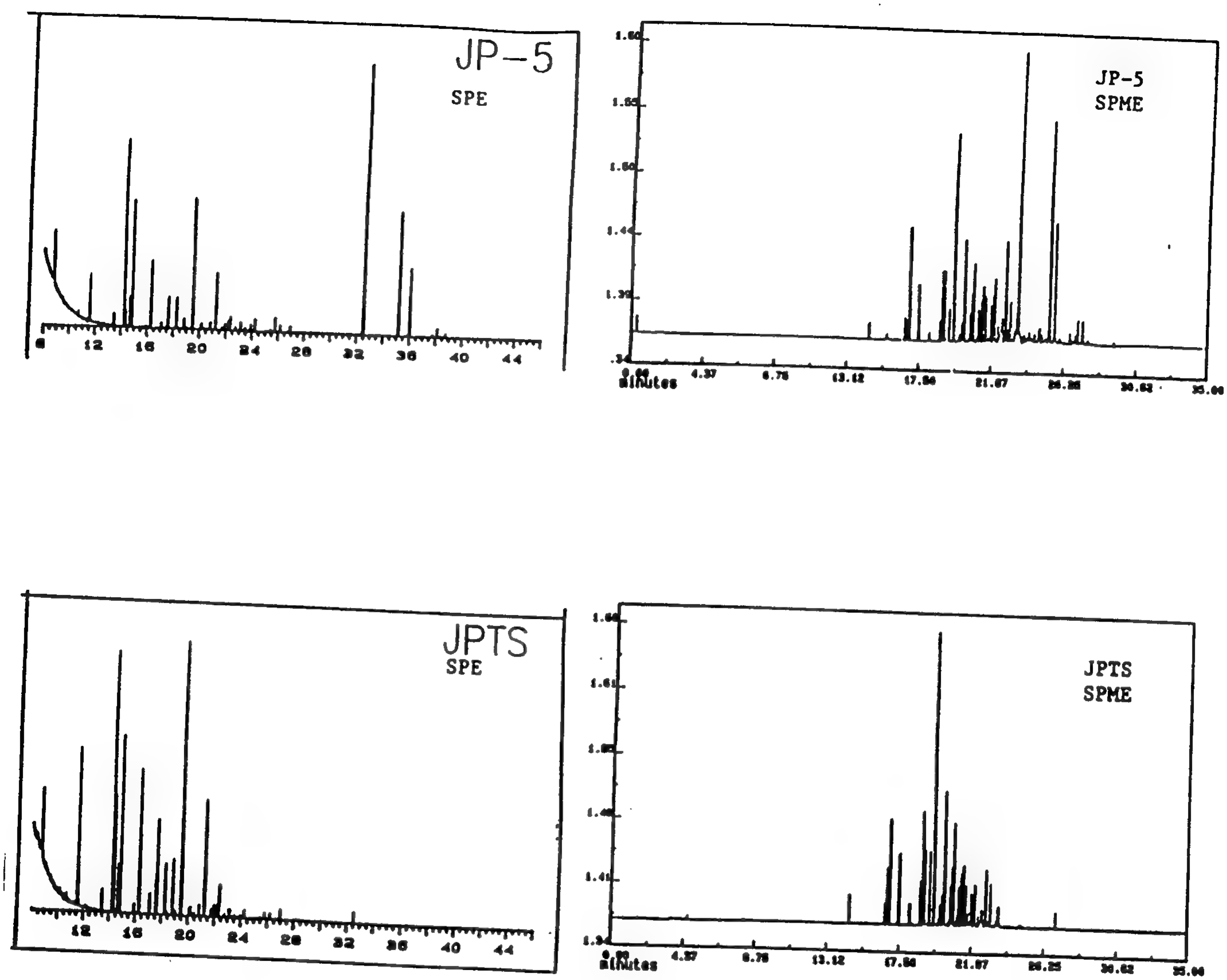

JP-4
SPE
8 12 16 20 24 28 32 36 40 44
JP-4
SPME
minutes
4.37
Jet-A
SPE
8 12 16 20 24 28 32 36 40 44
JET-A
SPME
minutes
4.37

Figure 7



JP-5
SPE
12 16 20 24 28 32 36 40 44
JP-5
SPME
minutes
4.37
13.12
17.50
21.87
30.62
35.00
JPTS
SPE
12 16 20 24 28 32 36 40 44
JPTS
SPME
minutes
4.37
13.12
21.87
26.25
30.62
35.00

References

1. Puyear, R.L.; Flecjenstein, K.J.; Montz, W.E.; Brammer, D. "Use of Reversed Phase C-18 Minicolumns for Concentrating Water Solvable Hydrocarbons," Bull. Contam. Toxicol., 1981, 95 790-797.

2. Mayfield, H.T.; Henley, M.V. "Classification of Jet Fuels Using High Resolution Gas Chromatography and Pattern Recognition," Monitoring Water in the 1990's: Meeting New Challenges, ASTM STP 1102, J.R. Hall and G.D. Glysson, Eds., American Society for Testing and Materials, Philadelphia, 1991.

3. Aruther, C.L.; Pawliszyn, J. "Solid Phase Microextraction with Thermal Desorption Using Fused Silica Optical Fibers," Anal. Chem., 1990, 62, 2145-2148.

4. Aruther, C.L.; Chai, M.; Pawliszyn, J. "Solvents Injection Technique for Microcolumn Separations," J.Microcol. Sepns., 1993, 5, 51-56.

5. Aruther, C.; Killam, L.; Buchholz, K.; Potter, D. Chai,M. Zhang,Z; Pawliszyn, J. "Solid Phase Microextraction: An Attractive Alternative," Envion. Lab., December/January, 1992/1993, pp.4-10

6. Aruther, C.L.; Killam, L.M.; Motlagh,S.; Lim, M.; Potter, D.W.; Pawliszyn, J. "Analysis of substituted Benzene Compounds in Groundwater Using Solid Phase Microextraction," Environ. Sci. Techn., 1992, 26, 979-983.

7. Aruther, C.L.; Pratt,K.; Motlagh, S.; Pawliszyn, J.; Belardi, R.P. "Environmental Analysis of Organic Compounds in Water Using Solid Phase Microextraction," J. H. Res. Chrom. Comm. 1992, 15, 741-744.

8. Hawthorne, S.B.; Miller, D.J.; Pawliszyn, J.; Aruther, C.L. "Solventless Determination of Caffeine in Beveridges Using Solid Phase Microextraction with Fused Silica Fibers," J. Chrom., 1992, 603, 185-191.

9. Burris, D.R.; MacIntyre, W.G. "Water Solubility Behavior of Binary Hydrocarbon Mixtures," Environmental Toxic. Chem., 1985, 4, 371-377.

10. Stomquist, V.K. "Solid Phase Microextraction as a Method for Quantifying Jet Fuel Contamination in Water," Final Report for Summer Graduate Research Program, Armstrong Laboratory, 1994.

11. Lavine, B.K.; Ritter,J; Faruque, A; Mayfield, H." Fuel Spill Identification by Solid Phase Microextraction and Pattern Recognition Techniques," Anal. Chim. Acta., Submitted

AN ANISOTROPIC MODELING STUDY OF THE MADE-2 SITE

Dale F. Rucker
Graduate Research Assistant

Department of Civil and Environmental Engineering
West Virginia University
Morgantown, WV 26506-6103

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Boiling Air Force Base, Washington, D.C.

August 1995

AN ANISOTROPIC MODELING STUDY OF THE MADE-2 SITE

Dale F. Rucker
Graduate Research Assistant

Department of Civil and Environmental Engineering
West Virginia University

Abstract

Previous efforts to replicate the transport of tritium along with four organic contaminants injected into the shallow alluvial aquifer of the MADE-2 site fell short of matching the observed phenomena. Discrepancies were blamed on improper head boundary conditions in the northern portion of the domain, showing a dipping of the head to the northwest. This did not logically follow the observations of the tritium plume, as it drifted to the northeast. Additional problems such as coding and an incomplete representation of the hydraulic conductivity matrix kept accurate results to a minimum. However, a new study is presented here, which tests the hypothesized existence of a meandering channel. The channel corresponds to high conductivity zones located in the mid-field of the domain. In order to model the effects of said channel, horizontal anisotropy was implemented into the flow model, which increased the velocity in the longitudinal direction.

As in prior modeling, MODFLOW was used to simulate the groundwater flow equation. Minor adjustments were made to the original source code to account for anisotropy varying from cell to cell, as opposed to layer by layer. The transport equation, MT3D, was applied to the conservative tracer and the organics, using the solution from the new flow model. Relatively good agreement between the observed and simulated plumes help substantiate the existence of an anisotropic media, supporting the existence of the meandering channel.

AN ANISOTROPIC MODELING STUDY OF THE MADE-2 SITE

Dale F. Rucker

INTRODUCTION

The fate and remediation of large jet fuel spills into the groundwater has been the topic of research by the Air Force for many years. Recently, the focus of investigation has been to study the physico-chemical attributes of hydrocarbon migration. The MADE (MAcroDispersion Experiment) experiment for example, conducted at the Columbus Air Force Base, Mississippi, tracked a conservative tracer along with four reactive aromatic organic compounds through an unconfined alluvial aquifer for approximately 15 months. A pulse injection of aqueous solution containing tritiated water, benzene, naphthalene, p-xylene, and o-dichlorobenzene was made into the saturated zone and the concentrations of each contaminant were recorded at specific time intervals using an array of multilevel samplers. The results of the sampling (Boggs et al.;1993) gave three-dimensional representations of the tracers, referred to as "snapshots".

In order to replicate the observed phenomena of the field conditions, a numerical model can be adapted and manipulated to fit particular field data. In turn a better understanding of complicated physical processes such as advection or dispersion can be achieved. Specifically, the MADE-2 experiment was modeled using public domain computer codes written in FORTRAN 77 to simulate groundwater flow and contaminant transport. MODFLOW, developed by McDonald and Harbaugh (1988), solved the groundwater flow equation for the piezometric head and specific discharge at discrete points within the domain. The solution from MODFLOW was then used as input to the transport code, MT3D (1990), written by Dr. Chunmiao Zheng, to solve the advective-dispersive-reactive equation for the concentration field. These codes were chosen for their reputation, ease of use, and applicability to the problem at hand.

Several attempts have been made to simulate the MADE-2 experiment, including works from Gray (1993), Gray and Rucker (1994), and Koch (1994). Concentrating mainly on tritium, which acts as the conservative tracer, the above studies fall short of matching the plume's tail in the far field. The

result: a stubby plume, extending no further than 150 meters downgradient, hence never reaching the 250 meters and beyond noticed with the observed data. Therefore, the objective for the 1995 Summer Research Program was to successfully simulate the tritium plume as well as the four organic contaminants by increasing advective transport, while retaining criteria placed upon head differences between that of the field conditions to simulated results. This work is an extension of previous efforts by the author and Dr. Donald Gray (an AFOSR Summer Faculty Fellow: Gray, 1992;1993, and Gray and Rucker, 1994).

FLOW MODELING

In order to properly estimate the head and velocity field in a groundwater modeling study, several aquifer parameters must first be identified, including specific storage, specific yield, porosity, and anisotropy. Specifically, conductivity anisotropy arises during the geological formation of the aquifer; sediments are deposited in such a way that permeability is higher in one direction than in others. The application of anisotropy to the MADE-2 site can be seen by the characterization of the aquifer. From Boggs et al.(1992) :

The shallow unconfined aquifer which immediately underlies the site consists of an alluvial terrace ... composed of poorly sorted to well sorted sandy gravel and gravely sand ... soil facies occur as irregular lenses and layers having horizontal dimensions ranging up to 8 m and vertical dimensions of less than 1 m.

The lenses described above indicate that the stratification of the soil particles may influence the groundwater flow in such a way that a horizontally anisotropic media is a better approximation than an isotropic one.

Modeling began by assuming the same aquifer conditions and discretization as that used in Gray and Rucker (1994). The only difference came with the addition of ten new hydraulic conductivity wells (borehole flowmeters), received in the Spring of 1995. Evaluation of the new data underwent the same evaluation as described in the previous study including interpolation and evaluation of each profile, horizontally, to obtain the horizontal conductivity and vertical leakance at each node of the computational grid

(Gray and Rucker; 1994). The averaged profiles were then log transformed, kriged using Geo-EAS Version 1.2.1 (Englund and Sparcs; 1991), and transformed back to unlogged values. This transformation was essential to avoid negative values in the kriging process. The values were written in MODFLOW format to the BCF2 module using the program BCF2MAKR, written by the author.

MODFLOW allows horizontal anisotropy in the block-centered flow package as long anisotropy is aligned along the principal axes of hydraulic conductivity. More exact, the horizontal anisotropy value is the ratio of column transmissivity (or hydraulic conductivity) to row transmissivity and one value may be given for each layer. A value of 1.0 is given for isotropic conditions.

Simulations began by first assuming the aquifer was isotropic. Table 1 lists the details of subsequent runs and accompanying results. The first simulation listed is from Gray and Rucker (1994) with runs M2-8-2 through M2-8-6 varying the TRPY from 1 to 100. This gave a wide range of results allowing for sensitivity in the data.

RUN #	TRPY	% DESCREP-ANCY	AVERAGE RMS (m)
M2-5-1	1	0.25	0.197
M2-8-2	1	3.84	0.2056
M2-8-3	2	3.70	0.2023
M2-8-4	10	2.90	0.2006
M2-8-5	50	1.40	0.2072
M2-8-6	100	0.84	0.2098

TABLE 1: Details of Simulations

The TRPY variable in the second column of Table 1 is the column to row anisotropic value given to all the layers. The last two values of 50 and 100 were used only to observe the head-gradient changes noticed from the high Darcy velocities. To date, no study has been found to indicate that such extreme conditions exist, therefore no physical meaning can be attached to the numbers.

The percent discrepancy column is the total volumetric difference, given in percent relative error between in flow and out flow of all sink/source terms as calculated by MODFLOW. This value accumulates throughout the entire simulation and can be found in the volumetric budget section. The large contrast noticed between last years' simulations of M2-5 to this years' can be

attributed to the fact that a Sun Sparc 2 was used for all previous modeling, compared to a 486/33MHz PC for this Summer's M2-8 simulations. Lastly, the average RMS column is the average of all the root mean square differences between averaged observed heads and simulated heads at ten separate observation locations in the MADE-2 site. The ten field piezometers, equipped with continuous water level recorders, each produced an RMS value, however only the average of all the differences is listed here. Since the differences were calculated using only ten locations, out of a possible 12474 cells (21 columns, 66 rows, 9 layers), it may be somewhat unclear to quantify the accuracy of the entire model with this lone RMS value. Yet no other method could be devised for the comparison. It does, on the other hand, follow the trend seen in Gray and Rucker (1994), giving an excellent comparison to last year's research.

Through careful examination of individual RMS values for each simulation, it was noticed that every observation site had a different "optimum" anisotropic value. For example, piezometer P53A had a smaller difference in head when anisotropy was equal to 1, whereas piezometer P60A was optimized at a value of 10. This prompted an additional simulation with anisotropy varying, not only by layer, but by column and row as well. Minor adjustments were made to the original MODFLOW source code in order to change the TRPY value being read from a one-dimensional array reader for all the layers, to a two-dimensional array reader, one array for each layer. All changes were saved in the new MODFLOW as MF95.FOR, and were demonstrated correct when tested against previous simulations. The coding changes can be found in the Appendix.

The new simulation, M2-8-7, used a smoothly heterogeneous anisotropic factor array for each of the nine layers, produced by kriging known values of anisotropy at the ten specified piezometer locations. The piezometers, being screened at two separate depths in the field, were divided into upper and lower and designated as A and B, respectively. The two levels correspond to model layers 1-4 and 5-9. The division seemed reasonable since the average depth of the upper piezometer is equivalent to layer 3 and the average depth of the lower piezometer compares to layer 8. All kriging was performed by Surfer Version 5 for Windows, a surface mapping and contouring package developed by Golden Software.

The results of this new model were predictable. The optimum anisotropic array produced the best averaged RMS to date, with a slightly lower value of 0.1944 m. However, the real test lies with the increased velocity field. Program VELCOMP compared the Darcy velocity arrays of the isotropic simulation to that of the varying anisotropic simulation and computed a relative percent increase at all discrete points, and an average of those increases. After comparing the longitudinal velocities, an average increase of 231.43% was calculated.

TRANSPORT MODELING

As mentioned, MT3D was the source code used to solve the advective-dispersive-reactive equation for the concentrations within the specified domain. The program has the same modular structure as MODFLOW and allows head and flux terms from the flow model to be used as input data. The version used for this study was V1.80. The code includes slight modifications to handle drying/rewetting of partially saturated cells, as described in Koch (1994).

MT3D was first applied to the tritium plume. Since tritium is nonreactive, i.e. no sorption onto the soil particles, and only undergoes radioactive decay (half-life of 12.26 years), this seemed to be the best place to start the test of the new effects of increased velocities. All parameters described in Gray and Rucker (1994) including porosity, diffusivity, and decay coefficient were used in this study. The only major modification was the increase of injection concentration from 0.0555 mCi/m^3 to 1000 mCi/m^3. The change was due to underflow errors occurring with the lower concentration. As long as all terms used in calculating the change in concentrations are linear, the above increase is acceptable because the results are presented for normalized concentrations. Dispersion coefficients of α_L=0.5 m, α_T=α_V=0.01 m, were used based on the advise of Dr. Koch.

The first simulation took approximately seven hours to complete on the Sun Sparc 10. Figure 1 shows the contours for a normalized distribution of concentrations at approximately 100 day intervals. The exact days were chosen to match the snapshots of the field data. The last day, day 456, does not have a corresponding snapshot for comparison and is shown here only to demonstrate the migration of tritium at the end of the simulation.

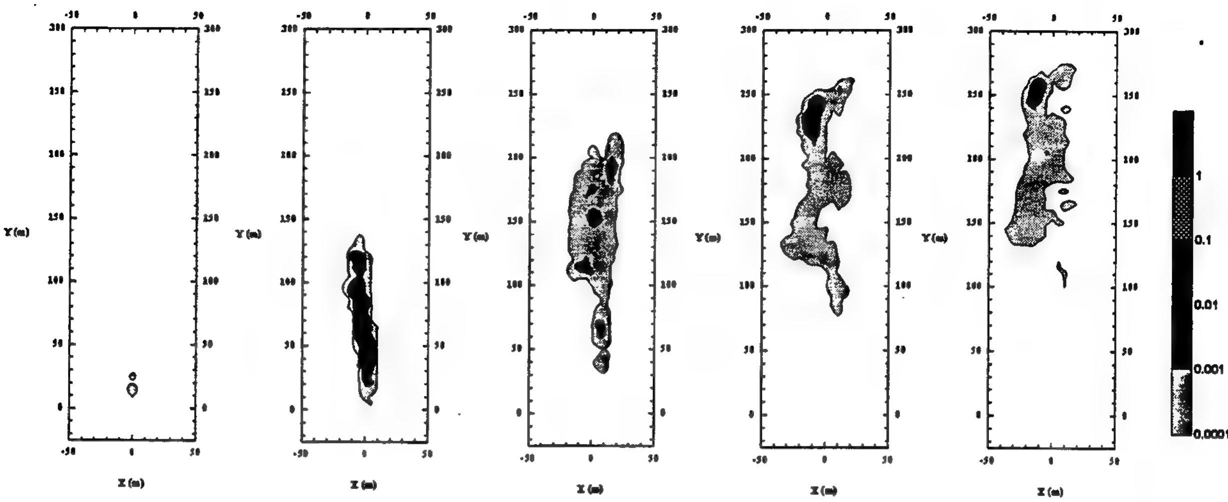


Figure 1: Normalized concentration (C/C_o) of tritium plume of Layer 4 for Run M2-8-7 for simulation days 44, 148, 240, 344, and 456. C_o is the injected concentration.

DISCUSSION

Through inspection, it was obvious that the predicted tritium plume was a poor fit to the observed data. Observed tritium contour plots of field data for days 44, 148, 240, and 344 are presented in the Appendix. The data was extracted at a level of 59.5 meters, which corresponds to model layer 4. The simulated plume movement in the longitudinal direction was much faster than observed, migrating entirely to the far field in the latter days. In reality the contaminant remains primarily in the near field with a tail of low concentrations extending to the mid and far regions. The observed phenomena is due to low hydraulic conductivity zones near the injection, with increasing conductivity in the mid field. In fact Rehfeldt et al. (1992) and Young (1995) describe the conductivity variation in the midfield as the result of a meandering river channel. Aerial photographs as well as sedimentological studies describing the depositional history of the alluvial aquifer were interpreted to indicate the existence of the channel. The channel may further explain the justification of anisotropy.

Inspection of hydraulic conductivity and transmissivity plots also concurred the location of the meandering channel. Figure 6 from Rehfeldt et al. (1992) shows the channel cutting through the mid section of the domain at an approximate 45° angle. Contour plots of the upper layers of the present investigation also show a pattern of higher conductivity in the same direction. Figure 2 shows the transmissivity of layer 3. The dashed line represents the direction in which the channel may have flowed.

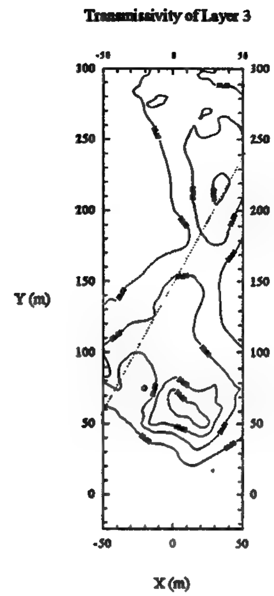


Figure 2: Transmissivity of Layer 3

Fueled by the findings of the previous set of experiments, a new MODFLOW simulation, M2-8-8, was performed. Rehfeldt et al. (1992) suggested that the presence of the hypothesized channel in the midfield may be in the regions of $K_h > 10^{-2}$ cm/s. Therefore, a high anisotropy value was assigned to the higher conductivity values, and low anisotropy to lower conductivity for run M2-8-8. Program TRPY tested the conductivity arrays for each layer and wrote the new anisotropy arrays to the block centered flow package, accordingly. After several tests, it was concluded that as conductivity increases above 3.0 m/d ($\sim 3.5 \times 10^{-3}$ cm/s), anisotropy would be set to 15, and lower values would have anisotropy equal to 1. Figure 3 shows a typical layer with the high anisotropy shaded in gray.

The flow model was rerun to obtain the new head and Darcy velocity values for transport simulation. The calculated averaged RMS head discrepancy for the simulation was 0.1958 m, slightly higher than M2-8-7, yet still better than simulation M2-5-1 from Gray and Rucker (1994). Head contours of Figure 4 also show that the new simulation to be a reasonable facsimile to the observed.

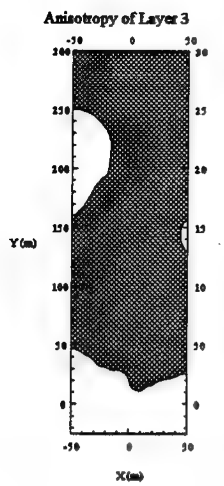


Figure 3: Anisotropy Array of Layer 3. Anisotropy=15 for shaded area.

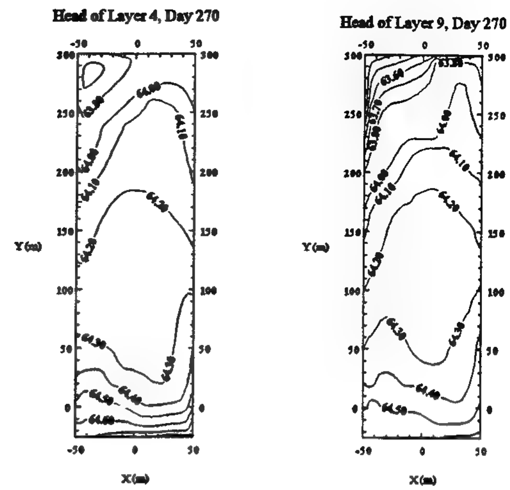


Figure 4: Head contours of Layers 4 and 9, Day 270 (March 8,1991).

Again, tritium was the first contaminant to be tested for solute migration. All other parameters as described above were adopted for this simulation as well. The resultant plume contours of Figure 5 show very good agreement between the observed and simulated concentrations. The contaminant moves slowly in the nearfield, increasing speed as it enters the mid regions of the

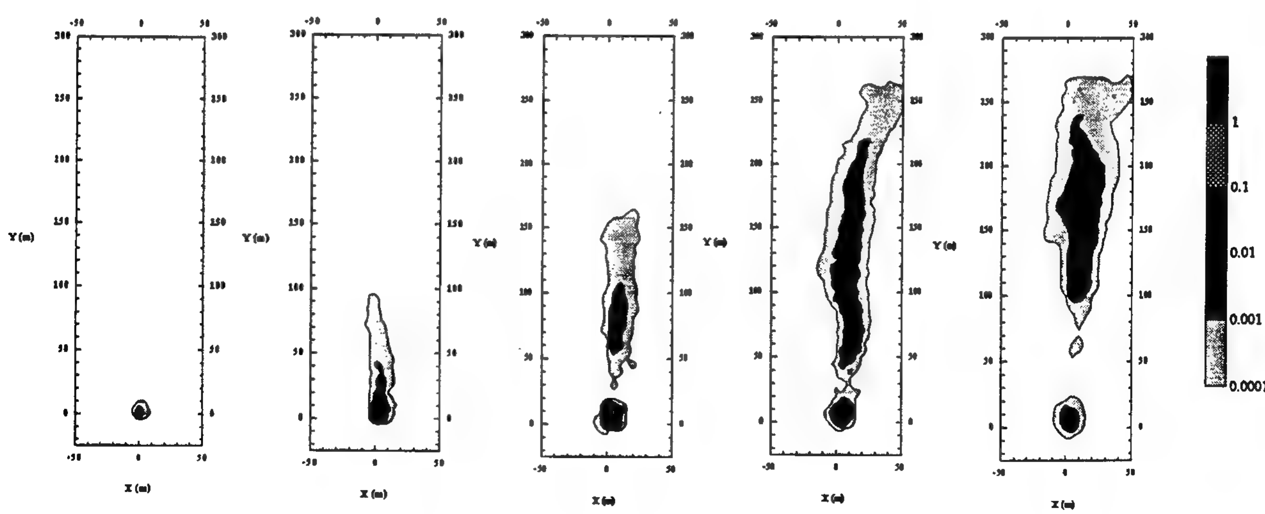


Figure 5: Normalized concentrations of tritium plume of Run M2-8-8 for simulation days 44, 148,240,344, and 456.

domain. The last days show a long tail trailing off to the northeast, just as observed.

The new flow model of M2-8-8 was also used to simulate the four organic plumes of benzene, naphthalene, o-dichlorobenzene, and p-xylene. With the exception of sorption and biodegradation, all parameters assumed in calculating tritium concentrations were used for the organics. MT3D simulates the effects of sorption and degradation in the chemical reaction (RCT) module. Sorption is implemented in the program through the use of retardation factor (R). A linear isotherm was assumed and the retardation factor was calculated as:

$$R_{i,j,k} = 1 + \frac{\rho_b}{\theta_{i,j,k}} K_d \tag{1}$$

where ρ_b is the bulk density (an aquifer parameter), K_d is the distribution coefficient, and $\theta_{i,j,k}$ is the porosity of cell i,j,k. Retardation factor is not entered directly in the transport model, but simulated by the above parameters. Boggs et al. (1993) obtained laboratory and field estimates of retardation. Knowing porosity, bulk density, and retardation, equation (1) was solved for K_d. Table 2 shows the values used to simulate sorption.

Organic	θ	Bulk Density x 10^6 ($^g/_m3$)	Retardation factor	K_d x 10^8	RC1 (d^{-1})	Initial Concentration ($^g/_m3$)
Benzene	0.32	1.77	1.30	5.42	0.0070	68.01
Naphthalene	0.32	1.77	1.42	7.59	0.0064	7.216
o-DCB	0.32	1.77	1.32	5.78	0.0046	32.75
p-Xylene	0.32	1.77	1.24	4.33	0.0107	41.44

TABLE 2: Transport parameters for organic transport simulations

Biodegradation was also simulated by using experimentally determined rate constants obtained from MacIntyre et al. (1993). Degradation kinetics were calculated from field data and were approximately first order. Table 2 lists

the rate constant as RC1, which is directly entered into the transport model without additional calculations.

The last column of Table 2 lists the initial concentrations used for transport simulations. These values were calculated by dividing the injected mass for each organic by the total volume of solution (9.7 m^3). The concentration is also entered directly in the transport model, within the basic transport module (BTN).

Each MT3D simulation ran for approximately 4-7 hours on a Sun Sparc 10. Contours were generated for every contaminant at snapshot intervals to compare with field results. Again, excellent agreement between observed and simulated plumes indicates that a channel within the midfield may exist, justifying an anisotropic media. However, in order to quantify the accuracy of the simulations, program MASSCALC was written to calculate the zeroth, first and second spatial moments, corresponding to the total mass, center of mass, and variance of the contaminant plumes, respectively.

Prior moment calculations on MADE-2 field data included three-dimensional numerical integration. The fundamental assumption behind integration is that concentration varies between descrete sampling points. However, for a finite difference grid, the concentration is constant over the entire cell. Therefore, MASSCALC calculated the moments by summation, allowing the concentration to remain constant within each cell. Tables 3(a) through 3(e) lists the results of MASSCALC from simulation M2-8-8.

TABLE 3(a): TRITIUM PLUME CHARACTERISTICS
<u>TIME</u>(simulation days)

	44	148	240	344	456
M/M_o	.9931	.9982	.9818	1.443	.9306
X_c (m)	1.06	1.97	4.88	9.367	13.25
Y_c (m)	1.61	6.17	27.06	93.38	159.51
Z_c (m)	57.12	58.36	59.42	59.61	58.85
σ_{11}^2 (m^2)	4.73	7.61	23.90	68.94	178.25
σ_{22}^2 (m^2)	16.93	92.61	1584	4081	3775
σ_{33}^2 (m^2)	.579	1.09	1.33	1.57	1.62

TABLE 3(b): BENZENE PLUME CHARACTERISTICS
TIME (simulation days)

	44	148	240	344	456
M/M_o	1.02	.477	.257	.150	.0369
X_c (m)	1.11	1.64	2.22	7.04	6.62
Y_c (m)	0.9	3.95	8.26	30.6	64.1
Z_c (m)	57.09	57.81	59.60	60.77	59.44
σ_{11}^2 (m^2)	4.69	6.34	9.19	15.1	21.4
σ_{22}^2 (m^2)	8.15	30.6	219	1490	4170
σ_{33}^2 (m^2)	.45	.852	1.15	1.24	.499

TABLE 3(c): NAPHTHALENE PLUME CHARACTERISTICS
TIME (simulation days)

	44	148	240	344	456
M/M_o	.977	.502	.309	.158	.0422
X_c (m)	1.09	1.50	2.79	4.56	3.67
Y_c (m)	.867	3.19	4.93	13.9	7.64
Z_c (m)	57.04	57.84	58.84	59.43	59.53
σ_{11}^2 (m^2)	4.57	5.61	6.91	7.49	4.85
σ_{22}^2 (m^2)	5.88	9.12	12.02	331.04	10.43
σ_{33}^2 (m^2)	.411	.651	1.19	1.27	.335

TABLE 3(d): p-XYLENE PLUME CHARACTERISTICS
TIME (simulation days)

	44	148	240	344	456
M/M_o	.78	.264	.095	.036	.006
X_c (m)	1.17	1.75	3.37	5.18	4.11
Y_c (m)	1.03	3.82	8.00	20.76	6.09
Z_c (m)	57.11	57.97	59.10	59.76	59.64
σ_{11}^2 (m^2)	4.86	6.46	7.43	10.70	3.65
σ_{22}^2 (m^2)	9.10	25.56	133.29	564.8	4.27
σ_{33}^2 (m^2)	.4767	.859	1.31	1.29	.229

TABLE 3(e): o-DCB PLUME CHARACTERISTICS
TIME (simulation days)

	44	148	240	344	456
M/M_o	1.12	.648	.1378	.0528	.0087
X_c (m)	1.17	1.63	3.37	5.18	4.11
Y_c (m)	1.03	3.85	8.00	20.76	6.09
Z_c (m)	57.11	57.91	59.10	59.76	59.64
σ_{11}^2 (m^2)	4.86	6.25	7.43	10.70	3.65
σ_{22}^2 (m^2)	9.10	26.54	133.29	564.8	4.27
σ_{33}^2 (m^2)	.4767	.799	1.31	1.29	.229

where M/M_o is total mass to injected mass, X_c, Y_c, and Z_c are the plumes' centroids in the respective directions, and σ_{11}^2, σ_{22}^2, and σ_{33}^2 are the principal second moments.

The tabulated results in the preceding five tables compare reasonably well to the observed field plume data, located in the Appendix. The largest discrepancy was noticed with longitudinal variance, σ_{22}^2. After rigorous inspection of the concentration data, the tail of the plumes seemed to separate from the bulk of the contaminant after entering the high velocity zone of the proposed channel. During the middle portion of the simulation, between days 148 and 344, the detached tail migrated farther downgradient, increasing the errors in the variance. The effects of the separation declined towards the end of the simulation due to biodegradation of the lowly concentrated tail.

One other unexplained oddity of the moment calculations is the ratio of total mass in the system to mass injected (M/M_o) for several of the contaminants. Benzene, for example, has an increase in total mass at 44 days of simulation. Since the MOC (Method of Characteristics) method, which does not conserve mass, was used to solve the advective portion of the transport equation, perhaps a two or as much as twelve percent over-estimation is possible. However, a value of 44% calculated for tritium at day 344 cannot be readily justified as over-estimation, nor can it be considered a fluke calculation. Several additional simulations were made to test the value, all giving the exact result of the original. Obviously, more time needs to be devoted to fully understand the problem. Later simulations will include a parametric

study, wherein one or two model parameters are changed by adding or subtracting a small tolerance.

CONCLUSIONS

1. Modeling of the MADE-2 site using an anisotropic media gave better approximation to the observed than any previous study using an isotropic media. This further proves the existence of a meandering channel through the site.

2. The additional hydraulic conductivity data increased the probability of accurately estimating the location of the channel. The locations of the flowmeters were approximately aligned in distinct rows at 45° angles from southwest to northeast. After kriging with Geo-EAS a definite pattern of channel deposits emerged.

3. The new simulated head distribution allowed tritium's tail to bend to the northeast, exactly as seen from field conditions. Perhaps the head boundary conditions, which did not change from Gray and Rucker(1994), should not be the focus of investigation and the real problem lies within the accelerated velocity needed to push the plume outward.

4. Simulation M2-8-8 did a relatively good job approximating the contaminant plumes. However, concrete proof is needed to justify the approach used to solve the problem. Pumping well data within the zone of the proposed channel would be very useful in calculating an exact anisotropic value. If the actual anisotropic value is determined to be a lower number, it would warrant further investigation on other aquifer parameters, such as porosity or hydraulic conductivity.

5. The MADE-2 experiment was a difficult problem to solve. Several years and hundreds of man hours were needed to finally understand the transport phenomena. However, without prior knowledge of the contaminant migration such as snapshot data, it would have been impossible to simulate as accurately as the previous study. Future studies that involve extremely heterogeneous aquifers should include all pertinent data including pumping well tests, more

exhaustive hydraulic conductivity data, and extensive research about the geologic history of the site.

ACKNOWLEDGMENTS

I would like to thank Dr. Donald Gray of West Virginia University; Dr. Tom Stauffer and Dr. Juminaou Yin of Armstrong Laboratory; and Dr. Manfred Koch of Temple University for their advise and assistance.

REFERENCES

J.M. Boggs, S.C. Young, L. Beard, L.W. Gelhar, K.H. Rehfeldt, and E.E. Adams, 1992. Field Study of Dispersion in a Heterogeneous Aquifer 1. Overview and Site Description, Water Resources Research, 28, 3281-3291.

J.M. Boggs, L.M. Beard, W.R. Waldrop, T.B. Stauffer, W.G. MacIntyre, and C.P. Antworth, 1993. Transport of tritium and Four Organic Compounds During a Natural-Gradient Experiment (MADE-2), EPRI Final Report TR-101998.

E. Englund and A. Sparcs, 1991. Geo-EAS 1.2.1 Geostatistical Environmental Software User's Guide, U.S. EPA 600/8-91/008.

D.D. Gray, 1993. Numerical Model of Groundwater Flow and Transport at the MADE-2 Site, U.S. Air Force Summer Faculty Research Program (SFRP) Reports, Volume 6, pp. 11-1 through 11-19, Air Force Office of Scientific Research, Bolling AFB, Washington, DC. NTIS ADA 278 693

D.D. Gray and D.F. Rucker, 1994. Improved Numerical Modeling of Groundwater Flow and Transport at the MADE-2 site. U.S. Air Force SFRP Reports, Volume 7, pp. 16-1 through 16-20, Air Force Office of Scientific Research, Bolling AFB, Washington, DC.

M. Koch, 1994. Application of the MT3D Solute transport Model to the MADE-2 Site. U.S. Air Force SFRP Reports, Volume 2,pp. 22-1 through 22-20, Air Force Office of Scientific Research, Bolling AFB, Washington, DC.

M.G. McDonald and A.W. Harbaugh, 1988. Techniques of Water-Resources Investigations of the U.S. Geological Survey, Chapter A1, A Modular Three-Dimensional Finite Difference Groundwater Flow Model, U.S. Government Printing Office, Washington DC.

W.G. MacIntyre, J.M. Boggs, C.P. Antworth, and T.B. Stauffer, 1993. Degradation Kinetics of Aromatic Organic Solutes Introduced Into a Heterogeneous Aquifer, Water Resources Research, 29, 4045-4051.

K.R. Rehfeldt, J.M. Boggs, and L.W. Gelhar, 1992. Field Study of Dispersion in a Heterogeneous Aquifer 3. Geostatistical Analysis of Hydraulic Conductivity, Water Resources Research, 28, 3309-3324.

S.C. Young, 1995. Characterization of High-K pathways by Borehole Flowmeter and Tracer Tests, Water Resources Research, 33, 311-318.

C. Zheng, 1990. MT3D, A Modular Three-Dimensional Transport Model for Simulation of Advection, Dispersion, and Chemical Reactions of Contaminants in Groundwater Systems, prepared for U.S. EPA Robert S. Kerr Environmental Research Laboratory, Ada, Oklahoma, by S.S. Papadopulos & Associates, Inc., Rockville, Maryland.

<u>APPENDIX</u>

Listing of changes to MODFLOW source code:

1) In the BCF2AL subroutine, space allocation needs to be larger.

Change:

```
      LCTRPY = ISUM
      ISUM = ISUM+NLAY
```

To:

```
      LCTRPY = ISUM
      ISUM = ISUM+ISIZ
```

2) In the BCF2RP subroutine,

Change:

```
      DIMENSION HNEW(NODES), SC1(NODES), HY(NODES), CR(NODES),
     1          CC(NODES), CV(NODES), ANAME(6,11), DELR(NCOL),
     1          DELC(NROW), BOT(NODES), TOP(NODES), SC2(NODES),
     1          TRPY(NLAY), IBOUND(NODES), WETDRY(NODES),
     1          CVWD(NODES)
```

To:

```
      DIMENSION HNEW(NODES), SC1(NODES), HY(NODES), CR(NODES),
     1          CC(NODES), CV(NODES), ANAME(6,11), DELR(NCOL),
     1          DELC(NROW), BOT(NODES), TOP(NODES), SC2(NODES),
     1          TRPY(NODES), IBOUND(NODES), WETDRY(NODES),
     1          CVWD(NODES)
```

and change:

```
      CALL U1DREL(TRPY,ANAME(1,8),NLAY,IN,IOUT)
      CALL U1DREL(DELR,ANAME(1,9),NCOL,IN,IOUT)
      CALL U1DREL(DELC,ANAME(1,10),NROW,IN,IOUT)
C
C2------READ ALL PARAMETERS FOR EACH LAYER
      KT=0
      KB=0
      DO 200 K=1,NLAY
      KK=K
C
C2A-----FIND ADDRESS OF EACH LAYER IN THREE DIMENSION ARRAYS.
      IF(LAYCON(K).EQ.1 .OR. LAYCON(K).EQ.3) KB=KB+1
      IF(LAYCON(K).EQ.2 .OR. LAYCON(K).EQ.3) KT=KT+1
```

```
      LOC=1+(K-1)*NIJ
      LOCB=1+(KB-1)*NIJ
      LOCT=1+(KT-1)*NIJ
C
C2B-----READ PRIMARY STORAGE COEFFICIENT INTO ARRAY SC1 IF TRANSIENT
      IF(ISS.EQ.0)
      CALL U2DREL(SC1(LOC),ANAME(1,1),NROW,NCOL,KK,IN,IOUT)
```

to:

```
      CALL U1DREL(DELR,ANAME(1,9),NCOL,IN,IOUT)
      CALL U1DREL(DELC,ANAME(1,10),NROW,IN,IOUT)
C
C2------READ ALL PARAMETERS FOR EACH LAYER
      KT=0
      KB=0
      DO 200 K=1,NLAY
      KK=K
C
C2A-----FIND ADDRESS OF EACH LAYER IN THREE DIMENSION ARRAYS.
      IF(LAYCON(K).EQ.1 .OR. LAYCON(K).EQ.3) KB=KB+1
      IF(LAYCON(K).EQ.2 .OR. LAYCON(K).EQ.3) KT=KT+1
      LOC=1+(K-1)*NIJ
      LOCB=1+(KB-1)*NIJ
      LOCT=1+(KT-1)*NIJ
C
C2B-----READ PRIMARY STORAGE COEFFICIENT INTO ARRAY SC1 IF TRANSIENT
      CALL U2DREL(TRPY(LOC),ANAME(1,8),NROW,NCOL,KK,IN,IOUT)
      IF(ISS.EQ.0)
      CALL U2DREL(SC1(LOC),ANAME(1,1),NROW,NCOL,KK,IN,IOUT)
```

*Note: The new TRPY variable is to be read by the U2DREL array reader before each SF1 array (if transient, or transmissivity if steady-state), as opposed to being read by the U1DREL array reader before DELR.

3) In BCF2FM, SBCF1C, SBCF2H, and SBCF2N, changes need to be made by changing the dimensions of TRPY (NLAY) to TRPY (NCOL, NROW,NLAY).

4) The last changes are to be made to the SBCF1C subroutine.
Change:

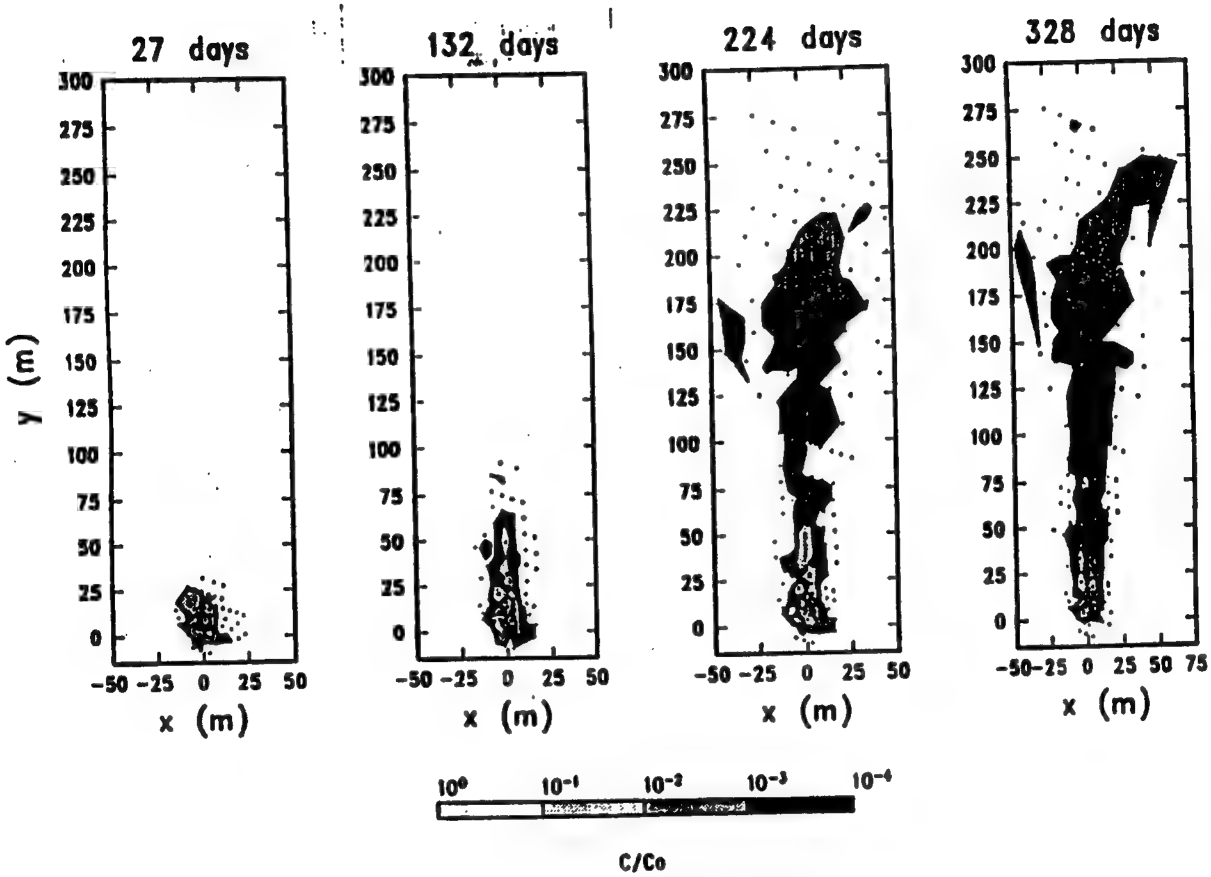
```
     YX=TRPY(K)
C
C1------FOR EACH CELL CALCULATE BRANCH CONDUCTANCES FROM THAT CELL
C1------TO THE ONE ON THE RIGHT AND THE ONE IN FRONT.
      DO 40 I=1,NROW
      DO 40 J=1,NCOL
      T1=CC(J,I,K)
```

to:

```
C    YX=TRPY(K)
C
C1------FOR EACH CELL CALCULATE BRANCH CONDUCTANCES FROM THAT CELL
C1------TO THE ONE ON THE RIGHT AND THE ONE IN FRONT.
```

```
DO 40 I=1,NROW
DO 40 J=1,NCOL
T1=CC(J,I,K)
YX = TRPY(J,I,k)
```

Observed Tritium Contour Plots:



Listing of observed plume characteristics:

TABLE 4(a): OBSERVED TRITIUM PLUME CHARACTERISTICS

	TIME				
	44	148	240	344	456
M/M_o	1.52	1.05	0.98	0.77	-
X_c (m)	0.0	-0.9	0.2	2.1	-
Y_c (m)	3.9	8.1	46.5	76.8	-
Z_c (m)	58.22	58.68	58.80	58.48	-
σ_{11}^2 (m^2)	10.3	94.4	4380	6560	-
σ_{22}^2 (m^2)	8.6	7.9	52.5	107	-
σ_{33}^2 (m^2)	2.0	1.3	2.5	2.9	-

TABLE 4(b): OBSERVED BENZENE PLUME CHARACTERISTICS

	TIME				
	44	148	240	344	456
M/M_o	0.92	0.43	0.23	0.07	0.06
X_c (m)	-0.2	-1.1	-0.8	-1.0	-1.1
Y_c (m)	3.8	6.3	12.4	7.7	7.9
Z_c (m)	58.13	58.68	59.24	58.90	58.67
σ_{11}^2 (m^2)	9.2	38.4	826	24.7	21.4
σ_{22}^2 (m^2)	7.9	6.9	6.1	8.4	10.7
σ_{33}^2 (m^2)	1.9	1.1	1.2	1.2	1.3

TABLE 4(c): OBSERVED NAPHTHALENE PLUME CHARACTERISTICS

	TIME				
	44	148	240	344	456
M/M_o	0.58	0.45	0.25	0.08	0.06
X_c (m)	-.3	-1.2	-1.4	-1.0	-1.8
Y_c (m)	3.5	5.8	6.6	7.2	7.3
Z_c (m)	58.05	58.54	59.03	58.71	58.62
σ_{11}^2 (m^2)	8.5	19.2	14.7	16.2	12.8
σ_{22}^2 (m^2)	7.3	6.7	6.1	4.8	4.7
σ_{33}^2 (m^2)	1.6	1.4	1.3	1.0	1.3

TABLE 4(d): OBSERVED p-XYLENE PLUME CHARACTERISTICS

	TIME				
	44	148	240	344	456
M/M_o	0.77	0.58	0.23	0.03	0.01
X_c (m)	-.2	-1.0	-1.2	-1.1	-1.3
Y_c (m)	3.6	6.1	6.4	10.5	6.4
Z_c (m)	58.04	58.60	58.93	58.66	57.93
σ_{11}^2 (m^2)	9.1	22.8	16.3	213	11.8
σ_{22}^2 (m^2)	7.2	5.8	6.4	4.6	4.0
σ_{33}^2 (m^2)	1.9	1.1	1.3	1.0	1.3

TABLE 4(e): OBSERVED o-DCB PLUME CHARACTERISTICS

	TIME				
	44	148	240	344	456
M/M_o	0.81	0.80	0.60	0.21	0.13
X_c (m)	0.0	-1.0	0.0	-0.2	-1.5
Y_c (m)	3.9	7.4	34.7	22.1	8.1
Z_c (m)	58.09	58.57	58.93	58.90	58.68
σ_{11}^2 (m^2)	11.3	64.4	3540	1180	21.9
σ_{22}^2 (m^2)	7.9	7.8	25.2	22.4	7.1
σ_{33}^2 (m^2)	2.1	1.4	2.0	1.6	1.8

APPLIED PERSONNEL TESTING: DIFFERENTIAL ITEM FUNCTIONING ON MEASURES OF COGNITIVE INFORMATION PROCESSING

Cynthia A. Searcy
Ph. D. Candidate
Department of Psychology

University of Georgia
Psychology Building
Athens, GA 30602

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1995

APPLIED PERSONNEL TESTING: DIFFERENTIAL ITEM FUNCTIONING ON MEASURES OF COGNITIVE INFORMATION PROCESSING

Cynthia A. Searcy
Ph. D. Candidate
Department of Psychology
Univsersity of Georgia

Abstract

A subset of the Applied Personnel Testing battery (APT) was examined with respect to differential item functioning (DIF). From the battery of 17 tests, 14 were assessed to determine the degree to which ability tests generated from a cognitive information processing perspective would be susceptible to nonconstant group performance, or differential responding for groups across ability levels. Four DIF indices were utilized: $Mh\chi^2$, DD, STD, and RMSWD. Negligible group differences in relative ability were observed for 10 of the 14 tests. Only moderate DIF was observed for two of the remaining four tests, likely due to the nature of the constructs being measured by the tests. The final two tests demonstrated some amount of DIF; it is likely that differential exposure to the concepts measured by these tests were a factor in these results.

APPLIED PERSONNEL TESTING:
DIFFERENTIAL ITEM FUNCTIONING ON
MEASURES OF COGNITIVE INFORMATION PROCESSING

Cynthia A. Searcy

Introduction

This study is part of a larger body of work designed to assess the psychometric properties and predictive validities of a battery of computerized cognitive ability tests collectively referred to as Advanced Personnel Testing (APT), designed from an information processing perspective. The primary purpose of the project is to gather an empirical knowledge base to transition these tests from experimental research to operational use. The goal of this particular study is to examine the set of tests with respect to differential item functioning (DIF), i.e. to determine whether items on these information processing tests function differently for subgroups in the population.

Concerns related to potential bias are central in test development. Historically, the term bias has been used in two semantically different contexts. The first relates to a statistical determination, when groups of equal ability have different probabilities of answering an item correctly. The second use, acquiring popularity in the 1960s and 1970s, refers to a social evaluation of fairness, where the differences between ethnic/racial groups in performance are due to an unfair capitalization on knowledge and skills more familiar to the white middle class, rather than actual differences in ability (Angoff, 1993). These divergent interpretations led to the creation of the term differential item functioning by Holland and Thayer (1988) to refer to the less value laden statistical evaluation of test items for nonconstant group differences in performance. The phrase item bias procedures refers to the entire process, encompassing both the statistical and the judgmental phases of evaluating test items, where the judgmental phase includes a logical analysis of the extent to which the instances of DIF appear to be invalid or unfair for individuals from a particular group (Camilli & Shepard, chap. 1).

Conducting item analyses is an important phase in test development. Identifying and removing items which are relatively more difficult for one group than another, where the difficulty is irrelevant to the test construct being measured, can improve the reliability of the test, as well as adding to the content and construct validity. DIF analyses work by detecting multidimensionality in a test, somewhat akin to conducting a factor analysis at the item level. As such, they can detect secondary factors which may or may not be a legitimate part of the test construct. Indications of DIF are therefore not necessarily indications of bias, and must be followed by a logical analysis of the items to determine the cause of differential validity and

whether those causes are relevant to the construct measured by the test (Camilli & Shepard, p. 21).

The APT Battery

The APT battery is a subset of tests derived from the Cognitive Abilities Measurement (CAM) battery (version 4), developed by researchers on the Lamp project (Learning Abilities Measurement Project) at the Armstrong Laboratory of the United States Air Force (Kyllonen, 1994). This larger test battery is based on a consensus information processing model used by a number of cognitive psychologists (Kyllonen, 1994). This model contains a number of features common among several well known models of information processing in the literature (e.g., Anderson, 1983; Langley, 1983). Included in this model are the following components: a short term working memory, two long-term memories (procedural and declarative memory), the perceptual processing system which interprets information from the environment and records that information in working memory, the cognitive processing system which links information from the two long-term memories to the short term working memory, and the motor processing system which is responsible for executing behavioral responses. These basic processes are assumed to be relatively autonomous yet operate together. Individual differences may be manifested in each of these autonomous components; the components are hypothesized to account for a substantial proportion of the individual differences observed in ability level.

The consensus information processing model offers some expectation for fairness in test items. Because the tests are designed to measure basic cognitive processes assumed not only to be relatively independent but also possessed by all individuals, these tests should be less susceptible to problems of differential item functioning due to differing cultural experiences. Stated another way, it is reasonable to expect that across all levels of ability, statistical indices of differential item functioning and the judgments of bias should be reduced in these tests when compared to more typical cognitive abilities tests.

Method

Sample. Participants (N=7088) were United States Air Force recruits in their 20th or 28th day of basic training. Participants were tested in groups of about 40. Participants were tested on Unisys 386 25 MHZ microcomputers with 15" multisync non-interlaced color monitors. Testing lasted approximately four hours with a 5-minute break half way through the testing session. Subtests were administered in two blocks. A fixed set of twelve tests were randomly administered in the first block, and the remaining five were randomly administered in the second block. Items were randomly presented within each subtest. Although approximately 93% of the participants finished the first block of tests, only approximately 70% completed the second block within time limits. Two of the tests in the second block, GKQ0 and GKS2, were used in this study and have substantially smaller sample sizes because of the failure to complete the tests within established time limits..

Of the 7088 participants, 5731 (80.9%) were males, 1353 (19.1%) were female, and 4 (.1%) had missing data on this variable. With respect to race/ethnicity, the composition was as follows: 5220 (73.6%) were Caucasian, 1060 (15.0%) were African American, 403 (5.7%) were Hispanic, 168 (2.4%) were Asian, 67 (.9%) were Native American, 166 (2.3%) other, and 4 (.1%) had missing data. Although the complete sample size was 7088, not all recruits finished every test; sample sizes for each test, overall as well as broken down by race/ethnicity and gender are located in Table 1.

Table 1. Sample Sizes for Each Group and Total Sample

	Gender		Race/Ethnicity			Total
Subtest	M	F	W	B	H	Sample
WMQ4	5676	1341	5184	1040	398	7017
WMS3	5686	1333	5187	1038	398	7019
WMV1	5662	1340	5182	1029	397	7002
INQ3	5660	1340	5174	1030	402	7000
INS2	5657	1338	5165	1032	401	6995
INV1	5678	1336	5178	1038	400	7014
FLQ2	5677	1341	5181	1040	400	7018
FLS1	5677	1346	5188	1041	402	7023
FLV3	5668	1340	5176	1039	399	7008
SLQ3	5675	1339	5175	1044	400	7014
SLS1	5675	1337	5185	1031	400	7012
SLV2	5673	1345	5189	1038	399	7018
GKQ0	4194	916	4063	526	262	5108
GKS2	4150	936	4036	517	274	5086

Cognitive Abilities Tests of APT Battery.

As described previously, the APT tests are a subset of the CAM4 battery, developed by researchers at the Armstrong Laboratories Human Resource Division by the LAMP team (see Kyllonen, 1994 for a detailed account of the theoretical development of the CAM4 battery). The Lamp tests were generated and categorized on the basis of the taxonomic definitions (see Table 2) which stemmed directly from the consensus model of information processing, as well as research which suggests that the content of the test also provides systematic variance in individual test scores. Although the taxonomy is multidimensional, the two major dimensions which drove test development include the cognitive processes and content domains.

The cognitive processes dimension includes six factors: capacity of working memory, execution speed of the three processing systems, declarative (fact) learning ability, procedural (skill) learning ability, breadth of declarative knowledge, and breadth of procedural knowledge. The second dimension, the content domain, consists of three general knowledge domain factors: verbal, quantitative and spatial. The three content domains are hypothesized to reflect individual differences in relative knowledge (e.g. verbal vs. quantitative) which are independent of the general differences in declarative or procedural knowledge

(Kyllonen, 1994).

Cognitive processes and content domains are hypothesized to be independent and are completely crossed in the taxonomy of tests, resulting in a 6 (cognitive processes) x 3 (content domain) matrix of cells. This resulting taxonomy provides a theoretical framework with which to evaluate existing tests and design new tests. A total of 59 tests have been developed to fit into 17 of the 18 cells of the CAM taxonomy (at this time no verbal general knowledge tests have been developed). Two to four tests have been created for each cell such that the multiple tests in each cell were generated by different test paradigms. For the Applied Personnel Testing (APT), one test was selected from each of the cells of the taxonomy, resulting in a total of 17 tests. The subtests were selected on the basis of their potential ability to predict success on a performance tests administered at the end of a training course subject to the condition that the paradigms used to generate the tests would not be duplicated within a row of the taxonomy; in other words, the tests designed to measure a particular cognitive process would be a function of different test paradigms for each of the content domains.

Table 2. CAM Factor Definitions

Cognitive Processes Factors	**Factor Definition**
Processing Speed	Speed of retrieving, transforming, or responding to stimuli.
Working Memory Capacity	The ability to process and store information simultaneously.
Declarative Learning	Ability to acquire new declarative knowledge (commit a novel fact to memory, then strengthen it with use, and minimuze decay over time).
Procedural Learning	Ability to acquire new rules and skill.
Declarative Knowledge	Depth, breadth, accessibility, and organization of declarative knowledge base.
Procedural Knowledge	Depth, breadth, accessibility, and organization of procedural knowledge base.
Knowledge Domain	
Verbal	Naturally verbal operations (e.g., linguistic transormations, category and synonymy judgments, part-of-speech classifications) performed on verbal stimuli (e.g., words, sentences).
Quantitative	Naturally quantitative operations (e.g., arithmetic operations, sign reversals, high-low and odd-even judgments) performed on quantitative stimuli (e.g., digits, numbers).
Spatial	Naturally spatial operations (e.g., rotation, reflection, physical matching, synthesizing) performed on spatial stimuli (e.g., Bruce figures, Palmer figures, partly filled matrices, Wechsler blocks).

Note. Table was adapted from Kyllonen (1994) CAM: A theoretical framework for cognitive abilities measurements. In D. K. Detterman (ed.), Current topics in human intelligence (Vol. 4) (pp. 307-359). Norwood, NJ: Ablex Publishing Corporation.

As such, the name of each test (e.g. WMQ4, Working memory quantitative) indicates the cognitive process and content domain being measured, and the number indicates the particular paradigm used to generate the test. For this particular study 14 of the 17 APT tests were selected for investigation: Working Memory-Quantitative (WMQ4), Working Memory Spatial (WMS3), Working Memory-Verbal (WMV1), Inductive

Reasoning-Quantitative (INQ3), Inductive Reasoning-Spatial (INS2), Inductive Reasoning-Verbal (INV1), Fact Learning-Quantitative (FLQ2), Fact Learning-Spatial (FLS1), Fact Learning-Verbal (FLV3), Skill Learning-Quantitative (SLQ3), Skill Learning-Spatial (SLS1), Skill Learning-Verbal (SLV2), General Knowledge-Quantitative (GKQ0) and General Knowledge-Spatial (GKS2). The three processing speed tests were not analyzed because rate scores are used to assess performance rather than number right. Detailed desriptions of these tests can be obtained elsewhere of the 17 Apt tests were used. (e.g., Christal, 1991; Kyllonen, 1994; Kyllonen et al., 1990).

Analytic Procedure

Description of DIF Indices. Four DIF indices were selected for the present study to determine if items function differently on the APT subtests for gender or ethnic groups after controlling for ability. All four are based on traditional item-analytic approaches. Three of the four have performed consistently in indicating DIF in previous studies (e.g. Linn et al., 1988). One additional index was selected on the basis of its ability to identify items with nonconstant or nonuniform DIF. Values for the four DIF indices were calculated for the 14 selected APT subtests for the following comparisons: males vs, females, Caucasians vs. African Americans, Caucasians vs. Hispanics, and a baseline white males vs. white males, obtained by randomly dividing the white male subgroup in two. The purpose of creating the final white male Reference and Focal groups was to contrast the magnitude of the DIF indices to the other comparison groups and to approximate the number of items identified as functioning differently by chance. All procedures were conducted in two phases. In the first phase, items reflecting potential DIF were identified and removed from the total test score. In the second phase, analyses were rerun. Each item removed in phase one was reintroduced into the total score for the analysis which concerned that particular item, and then removed from the remainder of the item analyses. Although the analyses for 11 of the 14 tests were based on the total test score, a slightly different approach was used for the three skill learning tests. The tests were administered in blocks, and scores improved across the blocks as subjects learned to perform the tasks. In an attempt to control for changing ability levels across blocks, DIF analyses were conducted on each block separately rather than the entire test; results were then combined to obtain averages for the test as a whole.

Mantel-Haenszel Chi-Square ($Mh\chi^2$). This chi-square statistic was proposed by Holland and Thayer (1986). The Mantel-Haenszel chi-square test is based on the formation of K 2 x 2 contingency tables for each item, where K is the number of observed total correct scores on the test. A contingency table is defined for each total test score at which one or more test takers in either group being compared has obtained. The reference and focal groups are identified, and items are scored one for correct and zero for incorrect. This index is distributed approximately as a chi-square statistic with a single degree of freedom, and is based on

the discrepancy of the observed number of correct responses on the item for a given total score on the test by the Reference group and the expected number. For an illustration of the computational definition of the Mantel-Haenszel Chi-Square based on the 2 x 2 contingency table see Figure 1 (Camilli & Shepard, 1994; Dorans & Holland, 1993).

		Score on the i[th] item		
		1	0	Total
	Reference	A_j	B_j	N_{rj}
Group	Focal	C_j	D_j	N_{fj}
	Total	M_{1j}	M_{0j}	T

$$MH\chi^2 = \frac{\left(\sum_{j=1}^{K} |A_j - E(A_j)| - .5\right)^2}{\sum_{j=1}^{K} var(A_j)}, \tag{1}$$

where

$$E(A_j) = \frac{N_{rj} M_{1j}}{T_j},$$

and

$$var(A_j) = \frac{N_{rj} N_{fj} M_{1j} M_{0j}}{T_j^2 (T_j - 1)}$$

<u>The Mantel-Haenszel Odds Ratio/Delta Difference (DD)</u>. The Mantel-Haenszel Odds Ratio is based on the same K 2 x 2 contingency tables as the $Mh\chi^2$. Equation 2 provides the definition for this index, referring again to the contingency table in Figure 1 (Camilli & Shepard, 1994; Dorans & Holland, 1993).

$$MHODDS = \frac{\sum_{j=1}^{K} A_j D_j / T_j}{\sum_{j=1}^{K} B_j C_j / T_j} \tag{2}$$

The MHODDS ratio can range from zero to infinity. By taking the simple natural logarithm transformation, the resulting statistic is symmetrical about zero, where zero has the interpretation of equal odds, thus improving the interpretability of the index. Positive values are interpreted as easier for the Reference group, and negative values are interpreted as easier for the Focal group. One further transformation places the index on the same metric historically used by ETS when examining DIF; applying equation 3 results in a rescaling of the log odds ratio to the delta metric, traditionally used by ETS (Camilli & Shepard, 1994; Dorans & Holland, 1993).

$$\text{Delta Difference} = (-2.35)\ln(\text{MHODDS}) \tag{3}$$

The Educational Testing Service uses the indices $Mh\chi^2$ and DD to classify items into categories A, B, and C, where the three categories indicate the degree of DIF in test items. The categories are defined as follows: A items have nonsignificant $Mh\chi^2$ or $|DD| < 1$, interpreted as demonstrating negligible DIF; B items have significant $Mh\chi^2$ and either (a) $|DD| < 1.5$ or (b) $|DD|$ not significantly different from 1; a B classification reflects moderate DIF, and items are examined for potential bias; C items have $|DD|$ significantly greater than 1.0 and $|DD| \geq 1.5$; these items reflect large DIF. Items with a C classification are typically removed or replaced (Camilli & Shepard, 1994; Dorans & Holland, 1993). For these analyses, the DD will be reported so that the preceding classification scheme can be used.

STD and RMWSD. The standardized p-difference (STD) and the root-mean-weighted squared difference (RMWSD) are indices which apply a weighting function to average differences (or squared differences) in difficulty across score levels, as seen in equations 4 and 5 (Camilli & Shepard, 1994; Dorans & Holland, 1993).

$$STD = \frac{\sum_{j=1}^{K} N_{Fj}(p_{Rj} - P_{Fj})}{\sum_{j=1}^{K} N_{Fj}}, \tag{4}$$

$$RMWSD = \sqrt{\frac{\sum_{j=1}^{K} N_{Fj}(p_{Rj} - p_{Fj})^2}{\sum_{j=1}^{K} N_{Fj}}}, \tag{5}$$

Although alternative methods of weighting the difference in difficulties for the Reference and Focal groups exist, by using the number of examinees at K in the focal group, the index is sensitive to the total scores most frequently obtained by Focal group examinees. This is useful, especially in situations where the Focal group sample is only moderate in size. These two indices can be interpreted as the average proportion difference controlling for the observed total score. STD ranges from -1 to +1, where positive values indicate that the item favors the Reference group, and negative items indicate that the item favors the Focal group. Values of STD between -.05 and +.05 (coded 1 in Table 13) are considered negligible; values between -.05 and -.10, or between +.05 and +.10 (coded 2 in Table 13) are moderate and should be inspected for effect; and values beyond -.10 or +.10 are atypical and need to be scrutinized (Dorans & Holland, 1993).

The primary difference between the STD and RMWSD is whether the difference in proportion correct is signed (unsquared) or squared. The advantage of the STD is that positive and negative differences in the proportions can cancel out; this is desirable when the differences are random, due to sampling errors and the discreteness of the total score distribution. The RMWSD, by squaring the differences in proportions, allows for the detection of nonuniform DIF. Although less research has been done on the RMWSD, it is useful when compared to the DD and Mhχ^2; if both the Mantel-Haenszel indices are small but the RMWSD is large, nonuniform DIF is suggested (Camilli & Shepard, 1994, chap. 5).

Results

Descriptive statistics based on the total sample are located in Table 3 and include the sample size, number of items, mean, standard deviation, skewness, kurtosis, and internal consistency reliability (Cronbach's coefficient alpha) for each of the 14 subtests. The scores are based on number correct composites, summed across items within a particular subtest. The number of items in the subtests ranged from 20 to 160. Sample sizes ranged from 6995 to 7023 for the first 12 tests, and 5086 (GKS2) to 5108 (GKQ0) for the remaining two. These sample sizes are based on the number of individuals from the total sample of 7088 to complete a particular subtest. Means for the subtests ranged from 8.92 (GKQ0) to 144.69 (SLQ3), and standard deviations ranged from 2.56 (INS2) to 18.45 (SLS1). Although means and standard deviations tended to be larger for subtests with more items, three tests were quite difficult: GKS2, SLS1, and FLS1 each had average number correct scores approximately 50 percent or less of the total possible score for that test. With respect to skewness and kurtosis, 12 of the 14 tests had negatively skewed distributions, and 10 of the 14 demonstrated positively kurtotic distributions. Internal consistency reliability estimates ranged from .32 to .95; GKQ0 (α=.32), GKS2(α=.50) INS2 (α=.62) and INV1 (α=.67) had the lowest estimates.

Means and standard deviations for each of the subgroups used in the DIF comparisons are located in

Table 3. APT Subtest Descriptive Statistics Based on Number Correct Scores for the Total Sample

Subtest	N	# Items	Mean	SD	Skewness	Kurtosis	Alpha
WMQ4	7017	32	24.31	5.75	-1.02	1.27	.85
WMS3	7019	27	17.45	6.40	-1.14	.84	.90
WMV1	7002	24	17.05	5.54	-1.22	1.04	.88
INQ3	7000	20	17.21	3.14	-2.12	6.15	.81
INS2	6995	20	16.01	2.56	-1.42	3.02	.62
INV1	7014	20	12.23	3.45	-.35	-.45	.67
FLQ2	7018	52	41.20	5.41	-.66	.63	.74
FLS1	7023	24	11.25	5.63	-.27	-.75	.87
FLV3	7008	32	26.69	5.19	-1.26	.96	.86
SLQ3	7014	160	144.69	17.85	-1.78	2.69	.86
SLS1	7012	96	48.14	18.45	.44	-.58	.94
SLV2	7018	72	62.63	11.68	-2.18	4.28	.95
GKQ0	5108	28	8.92	2.74	.13	-.13	.32
GKS2	5086	40	34.53	2.87	-.92	1.73	.50

Table 4. Average total subtest scores tended to be larger for males than females. Mean scores were higher on all subtests for Caucasians than both Hispanics and African Americans, and Hispanics tended to score higher than African Americans. Internal consistency reliability estimates for the subgroups are located in Table 6. The estimates of reliability are stable across the groups for both gender and race/ethnicity, with the exception of GKQ0. The reliability of the items for women (α=.24) is consistently lower than for any other subgroup examined, where the reliability remains above α=.30.

Summary results from the DIF analyses are located in Table 6. Included are means and standard deviations of the four DIF indices, Mhχ^2, DD, STD, and RMWSD, for the comparisons of gender and race/ethnicity, where Caucasians were compared to both African Americans and Hispanics. Table 7 contains the number of items from each comparison falling into the categories A, B, and C as used by ETS. Table 8 contains ratings of items in terms of negligible, moderate, and severe DIF based on STD, coded 1 (negligible DIF), 2 (moderate DIF) and 3 (severe DIF). Finally, Table 9 contains information conveying the direction of DIF for items flagged as B or C using DD. In general, the Mhχ^2 values associated with the gender analyses were larger than the other comparison groups, and values associated with the white male comparison analyses were smallest. The average differences in the magnitude of the chi-square values could in part be accounted for by the differences in sample sizes since the magnitude of the chi-square tends to increase as sample size increases (see e.g., Linn et al., 1988; Welsh, Androlewicz, & Curan, 1990). This is a likely explanation for the difference in the chi-square values for the Caucasian vs. Hispanic analyses as compared with the other Reference and Focal groups. However, the fact that the female sample is comparable in size to the African American sample, yet yields substantially larger chi-square values, suggests that more accounts for the

Table 4. Means and Standard Deviations (in parentheses) by Group and Total Sample

	Gender		Race/Ethnicity			Total
Subtest	M	F	W	B	H	Sample
WMQ4	23.81	23.61	24.74	22.22	23.94	24.31
	(5.44)	(5.72)	(25.49)	(6.37)	(6.12)	(5.75)
WMS3	17.85	15.74	18.13	14.19	17.32	17.45
	(6.26)	(6.69)	(6.02)	(7.25)	(6.21)	(6.40)
WMV1	17.01	17.21	17.66	14.83	16.14	17.05
	(5.60)	(5.32)	(5.18)	(6.21)	(5.85)	(5.54)
INQ3	17.32	16.74	17.31	16.75	16.92	17.21
	(3.12)	(3.19)	(3.11)	(3.03)	(3.27)	(3.14)
INS2	16.15	15.40	16.20	15.15	15.76	16.01
	(2.48	(2.80)	(2.42)	(2.93)	(2.70)	(2.56)
INV1	12.31	12.07	12.63	10.80	11.61	12.23
	(3.47)	(3.33)	(3.39)	(3.34)	(3.52)	(3.45)
FLQ2	41.09	41.67	41.30	40.75	41.13	41.20
	(5.45)	(5.20)	(5.40)	(5.40)	(5.50)	(5.41)
FLS1	11.62	9.70	11.82	8.60	10.94	11.25
	(5.65)	(5.28)	(5.51)	(5.66)	(5.31)	(5.63)
FLV3	26.71	26.60	26.97	25.34	26.73	26.69
	(5.18)	(5.22)	(5.04)	(5.58)	(5.14)	(5.19)
SLQ3	144.96	143.53	146.15	137.25	144.36	144.69
	(17.75)	(18.24)	(16.83)	(21.22)	(17.12)	(17.85)
SLS1	49.68	41.64	49.75	40.19	47.28	48.14
	(18.76)	(15.67)	(18.53)	(15.69)	(18.52)	(18.45)
SLV2	62.84	61.73	63.34	59.30	62.73	62.63
	(11.64)	(11.83)	(10.90)	(13.97)	(12.35)	(11.68)
GKQ0	9.16	7.81	9.07	8.10	8.46	8.90
	(2.72)	(2.57)	(2.74)	(2.64)	(2.74)	(2.74)
GKS2	34.71	33.73	34.75	33.29	34.00	34.53
	(2.79)	(3.07)	(2.78)	(3.14)	(2.33)	(2.87)

differences than mere sample size discrepancies. A second interesting comparison lies between the unsigned index RMWSD and the signed index STD. The RMWSD, useful for detecting nonuniform DIF, in all cases was substantially larger than the STD. To the extent that random error causes positive and negative differences between groups in terms of mean responses across ability levels, this index will be inflated. The similarity across all the subtests of the RMWSD for the various Reference and Focal groups with the baseline white male comparison indicates that the nonconstant differences between groups across ability levels for the items is somewhat trivial. An examination of item response curves for items resulting in moderate to high DIF indices suggest that most of the nonuniformity in curves is a function of the lower ability levels, where guessing and the negatively skewed distributions are factors. Guessing is a factor which likely affects results, especially for the subtests where one of two response options is correct, allowing 50 percent correct on the subtest by chance alone. Further, negatively skewed distributions are a function of proportionately

Table 5. Coefficient Alpha Internal Consistency Estimates of Reliability by Group and Total Sample

	Gender		Race/Ethnicity			Total
Subtest	M	F	W	B	H	Sample
WMQ4	.85	.86	.84	.86	.87	.85
WMS3	.90	.90	.89	.92	.89	.90
WMV1	.88	.87	.86	.89	.87	.88
INQ3	.81	.79	.81	.80	.81	.81
INS2	.61	.65	.59	.67	.65	.62
INV1	.67	.65	.66	.63	.67	.67
FLQ2	.74	.73	.73	.70	.75	.74
FLS1	.87	.86	.89	.84	.85	.87
FLV3	.86	.86	.86	.86	.86	.86
SLQ3	.96	.96	.96	.96	.96	.96
SLS1	.94	.92	.94	.92	.94	.94
SLV2	.96	.95	.95	.96	.96	.95
GKQ0	.31	.24	.32	.30	.33	.32
GKS2	.48	.51	.48	.51	.53	.50

fewer participant scores occurring at lower ends of ability, resulting in less stable estimates of performance at those lower levels.

WMQ4. Mean chi-square values for the 32 item subtests ranged from .58 (white male comparison group) to 2.69 (gender). DD ranged from -.03 (Caucasian vs. Hispanic) to .06 (gender). Although the RMWSD in all comparisons was substantially larger than the unsigned STD, the similarity in values for the gender and white male comparison suggest that differences in item responses across ability levels were negligible. Using the ETS grading system, only one item was rated B, demonstrating moderate DIF for the Caucasian vs. Hispanic comparison; this difference favored the Caucasian Reference group (see Tables 12 and 14). The alternative grading system based on the magnitude of STD indicated that one item each the gender and Caucasian vs. Hispanic comparisons resulted in moderate DIF (see Table 13). Other than these three instances, the item analyses, for all comparisons, revealed negligible DIF for this subtest.

WMS3. Mean chi-square values for this subtest ranged from .08 (white male comparison) to 3.34 (gender). Mean DD values ranged from -.04, favoring the Reference group in the Caucasian vs. African American comparison to .07, favoring females in the gender comparison. Values of RMWSD were larger than STD, but results from the gender and ethnic comparisons were not substantially different from the white

male comparison. Using the ETS grading system, as well as that based on the STD, all of the items in each of the comparisons resulted in negligible DIF.

WMV1. Average chi-square values ranged from 1.09 for the white male comparison to 2.86 for the gender comparison. Mean DD values ranged from -.03 to .04. Again, the low value corresponded to the Caucasian vs. Hispanic comparison favoring the Reference group, and the higher value corresponded to the gender comparison favoring the Focal group. No items were graded B or C based on the $Mh\chi^2$ and DD. Only one item was graded moderate based on STD values; this item was flagged in the Caucasian vs. Hispanic analyses.

INQ3. Chi-square values ranged from .37 for the white male comparison to 3.55 for the gender comparison. DD ranged from .02 to .12. Although the gender analyses resulted in the low average DD and the Caucasian vs. Hispanic analyses resulted in the large average DD indices, it should be noted that on average, for this tests differences across ability levels tended to favor the Focal groups. Only one item received a grade more severe than A (for Caucasians vs. Hispanics): one item received a grade of C, favoring the Focal group of Hispanics.

INS2. Mean chi-square values ranged from 1.02 for the analyses including Caucasian vs. Hispanic comparison to 10.92 for the gender analyses. Chi-square values were moderately large for this test in the gender analyses when compared to many of the other subtests. Mean values of DD ranged from -.14, again for the Caucasian vs. Hispanic analyses, to .06 for the white male comparison analyses. One item was flagged from the Caucasian vs. Hispanic analyses and given the grade of C; this item favored the Caucasian Reference group.

INV1. Chi-square values were substantially larger for this subtest, especially for the gender analyses. Mean chi-square values ranged from 1.56 for the white male comparison to 59.89 for gender. Whereas most of the chi-square values did not differ substantially across the various Reference and Focal comparisons for most other subtests, they differed greatly for this subtest. Although the chi-square values were much greater, on average, the differences between Reference and Focal groups did not inordinately favor either the Reference or Focal group for any comparisons. Mean values of DD ranged from -.00 (Caucasian vs. African American comparison) to .04 (white male comparison). Two items received grades other than A in the gender analyses; the item afforded a B favored the Reference group, and the item afforded a C favored the Focal group. Although differences tended to average out, the results should be interpreted with caution. Because of the process of removing biased items from the analyses after the first run, only nine items were used to block individuals in the gender comparison. It is not known the extent to which less than half of the items on a subtests are representative of the total set of items. Based on the STD categorization scheme,

Table 6. Means and Standard Deviations (in parentheses) of DIF Indices for the Subtests

Subtest	M-F				W-B				W-H				W-W			
	$Mh\chi^2$	DD	STD	RMWSD	$Mh\chi^2$	DD	STD	RMWSD	$Mh\chi^2$	DD	STD	RMWSD	$Mh\chi^2$	DD	STD	RMWSD
WMQ4	2.69	.06	.00	.25	1.38	.02	.00	.28	1.42	-.03	.00	.33	-.00	-.00	.00	.26
	(4.46)	(.33)	(.02)	(.02)	(2.32)	(.26)	(.02)	(.02)	(2.17)	(.45)	(.02)	(.02)	(.85)	(.16)	(.01)	(.02)
WMS3	3.34	.07	.00	.25	1.64	-.04	.00	.26	.80	.00	.00	.32	.08	.01	.00	.24
	(4.10)	(.33)	(.02)	(.02)	(2.35)	(.30)	(.02)	(.02)	(1.13)	(.30)	(.02)	(.03)	(.20)	(.20)	(.01)	(.02)
WMV1	2.86	.04	.00	.26	1.56	.02	.00	.27	1.12	-.03	.00	.32	1.09	-.02	.00	.24
	(3.83)	(.29)	(.02)	(.02)	(2.13)	(.25)	(.02)	(.02)	(1.62)	(.35)	(.02)	(.03)	(1.50)	(.20)	(.01)	(.02)
INQ3	3.55	.02	.00	.22	2.62	.02	.00	.23	2.02	.12	.00	.27	.37	.03	.00	.21
	(8.70)	(.29)	(.02)	(.02)	(4.03)	(.31)	(.01)	(.03)	(2.94)	(.53)	(.02)	(.03)	(.40)	(.19)	(.00)	(.02)
INS2	10.92	.04	.00	.23	2.91	.03	.00	.24	1.02	-.14	.00	.29	1.28	.06	.00	.22
	(12.81)	(.42)	(.02)	(.03)	(3.92)	(.34)	(.01)	(.02)	(1.57)	(.56)	(.01)	(.03)	(2.03)	(.26)	(.01)	(.02)
INV1	59.89	.02	.00	.25	22.59	-.00	.00	.26	2.78	.02	.00	.31	1.56	.04	.00	.23
	(96.07)	(.60)	(.05)	(.05)	(21.69)	(.45)	(.04)	(.03)	(3.15)	(.48)	(.04)	(.03)	(1.40)	(.20)	(.02)	(.03)
FLQ2	6.36	.03	.00	.27	2.75	.03	.00	.28	1.01	.02	.00	.33	.43	-.01	.00	.27
	(10.76)	(.43)	(.03)	(.03)	(4.62)	(.38)	(.02)	(.03)	(1.44)	(.45)	(.02)	(.03)	(.62)	(.15)	(.01)	(.02)
FLS1	4.22	.01	.00	.24	2.54	.03	.00	.25	.96	.04	.00	.31	.92	.04	.00	.23
	(4.74)	(.39)	(.02)	(.02)	(3.37)	(.36)	(.02)	(.02)	(1.25)	(.33)	(.02)	(.03)	(1.25)	(.19)	(.01)	(.02)
FLV3	2.72	-.00	.00	.25	1.73	-.03	.00	.28	1.10	-.03	.00	.32	1.12	-.04	.00	.24
	(3.18)	(.33)	(.02)	(.02)	(2.36)	(.30)	(.02)	(.03)	(1.95)	(.40)	(.02)	(.03)	(1.21)	(.25)	(.01)	(.02)
SLQ3	2.93	.10	.00	.25	1.13	-.00	.00	.27	.99	-.00	.00	.30	1.13	.01	.00	.25
	(3.89)	(.47)	(.02)	(.02)	(1.68)	(.35)	(.02)	(.02)	(1.43)	(.54)	(.02)	(.03)	(1.66)	(.36)	(.02)	(.02)
SLS1	5.00	-.06	.00	.25	1.59	.01	.00	.26	.72	.00	.00	.31	1.09	.00	.00	.24
	(8.33)	(.29)	(.02)	(.02)	(1.95)	(.22)	(.02)	(.02)	(.97)	(.26)	(.02)	(.03)	(1.46)	(.18)	(.02)	(.02)
SLV2	2.75	.03	.00	.24	2.04	.03	.00	.26	1.30	.02	.00	.31	1.12	.01	.00	.23
	(5.21)	(.42)	(.01)	(.02)	(4.03)	(.37)	(.02)	(.02)	(1.88)	(.55)	(.02)	(.03)	(1.73)	(.30)	(.01)	(.03)
GKQ0	102.41	-.16	.00	.27	5.57	.05	.00	.28	1.64	.01	.00	.30	1.03	.00	.00	.23
	(17.16)	(.96)	(.06)	(.05)	(8.21)	(.51)	(.04)	(.03)	(2.68)	(.49)	(.04)	(.03)	(.98)	(.21)	(.02)	(.03)
GKS2	4.48	-.08	.00	.27	.95	.07	.00	.30	.96	.06	.00	.30	.81	-.04	.00	.23
	(7.48)	(.45)	(.02)	(.04)	(1.41)	(.44)	(.02)	(.04)	(1.41)	(.71)	(.02)	(.05)	(.89)	(.38)	(.01)	(.03)

Table 7. Distribution of ETS DIF classifications across Subtests

	M-F			W-B			W-H			W-W		
Subtest	A	B	C	A	B	C	A	B	C	A	B	C
WMQ4	32	0	0	32	0	0	31	1	0	32	0	0
WMS3	27	0	0	27	0	0	27	0	0	27	0	0
WMV1	24	0	0	24	0	0	24	0	0	24	0	0
INQ3	20	0	0	20	0	0	19	0	1	20	0	0
INS2	20	0	0	20	0	0	19	0	1	20	0	0
INV1	18	1	1	20	0	0	20	0	0	20	0	0
FLQ2	51	1	0	51	1	0	52	0	0	52	0	0
FLS1	24	0	0	24	0	0	24	0	0	24	0	0
FLV3	32	0	0	32	0	0	32	0	0	32	0	0
SLQ3	157	3	0	160	0	0	155	5	0	158	2	0
SLS1	96	0	0	96	0	0	96	0	0	96	0	0
SLV2	71	0	1	71	1	0	70	2	0	71	1	0
GKQ0	22	3	3	27	1	0	26	2	0	28	0	0
GKS2	40	0	0	40	0	0	39	0	1	40	0	0

Table 8. Distribution of STD DIF classifications across Subtests

	M-F			W-B			W-H			W-W		
Subtest	1	2	3	1	2	3	1	2	3	1	2	3
WMQ4	31	1	0	32	0	0	31	1	0	32	0	0
WMS3	27	0	0	27	0	0	27	0	0	27	0	0
WMV1	24	0	0	24	0	0	23	1	0	26	0	0
INQ3	20	0	0	20	0	0	20	0	0	20	0	0
INS2	20	0	0	20	0	0	20	0	0	20	0	0
INV1	16	3	1	18	2	0	18	2	0	20	0	0
FLQ2	49	3	0	51	1	0	51	1	0	52	0	0
FLS1	24	0	0	24	0	0	24	0	0	24	0	0
FLV3	24	0	0	24	0	0	23	1	0	24	0	0
SLQ3	160	0	0	160	0	0	159	1	0	160	0	0
SLS1	94	2	0	96	0	0	95	1	0	96	0	0
SLV2	72	0	0	71	1	0	72	0	0	72	0	0
GKQ0	24	2	2	24	4	0	26	1	1	28	0	0
GKS2	38	2	0	40	0	0	40	0	0	40	0	0

Table 9. Distribution of STD DIF classifications across Subtests

	M-F			W-B			W-H			W-W		
Subtest	R>F	R=F	R<F	R>F	R=F	R<F	R>F	R=F	R<F	R>F	R=F	R<F
WMQ4	0	32	0	0	32	0	1	31	0	0	32	0
WMS3	0	27	0	0	27	0	0	27	0	0	27	0
WMV1	0	24	0	0	24	0	0	24	0	0	24	0
INQ3	0	20	0	0	20	0	0	19	1	0	20	0
INS2	0	20	0	0	20	0	1	19	0	0	20	0
INV1	1	18	1	0	20	0	0	20	0	0	20	0
FLQ2	0	51	1	0	51	1	0	52	0	0	52	0
FLS1	0	24	0	0	24	0	0	24	0	0	24	0
FLV3	0	32	0	0	32	0	0	32	0	0	32	0
SLQ3	0	157	3	0	160	0	2	155	3	1	158	1
SLS1	0	96	0	0	96	0	0	96	0	0	96	0
SLV2	0	71	1	1	71	0	0	70	2	0	71	1
GKQ0	4	22	2	1	27	0	1	26	1	0	28	0
GKS2	0	40	0	0	40	0	1	39	0	0	40	0

three items resulted in moderate DIF and one more substantial DIF for the gender analyses; two items from the Caucasian vs. African American comparison resulted in moderate DIF; finally, two items from the Caucasian vs. Hispanic comparison resulted in moderate DIF.

FLQ2. Mean chi-square values ranged from .92 (white male comparison) to 6.36 (gender comparison). Average magnitudes of differences were quite small, ranging from -.01 (white male comparison) to .03 (gender and Caucasian vs. African American comparisons). One item from both the gender and Caucasian vs. African American analyses was graded B using the ETS grading system; in both cases the differences favored the Focal group. Based on the STD, three items from the gender analyses were rated moderate in terms of DIF, and one item each from the Caucasian vs. African American and Caucasian vs. Hispanic analyses was rated moderate.

FLS1. Average chi-square values ranged from .92 to 4.74, corresponding to the white male and gender analyses. Mean DD values were small and positive, ranging from .01 to .04. All items demonstrated negligible DIF on this subtest.

FLV3. Chi-square averages on this subtest ranged from 1.10 (Caucasian vs. Hispanic) to 2.72 (gender), and mean DD values ranged from -.04 (white male) to -.00 (gender). No items were graded B or C based on the ETS system, and one item from the Caucasian vs. Hispanic analyses was rated moderate in DIF.

SLQ3. For this subtest, average chi-square values ranged from .99 to 2.93, again for the white male and gender comparison groups. Average DD ranged from -.00 (Caucasian vs. African American and Caucasian vs. Hispanic analyses) to .10 for gender. Three items from the gender analyses, five items from the Caucasian vs. Hispanic analyses, and two items from the white male analyses were rated B on the basis of

$Mh\chi^2$ and DD. All three items from the gender analyses, three items from the Caucasian vs. Hispanic, and one item from the white male analyses favored the focal group. Two items from the Caucasian vs. Hispanic and one item from the white male comparison favored the Reference group.

SLS1. Mean chi-square values ranged from .72 (Caucasian vs. Hispanic) to 5.00 (gender), and mean DD values ranged from .00 (Caucasian vs. Hispanics and white males) to -.06 (gender). No items received higher than an A rating based on $Mh\chi^2$ and DD. Two items from the gender analyses were rated moderate in DIF with respect to STD, and one item was rated moderate from the Caucasian vs. Hispanic analyses.

SLV2. As with the majority of the subtests, the lowest average chi-square resulted from the white male analyses, and the highest resulted from the gender analyses. The magnitude of the DIF was quite small, as evidenced by mean DD values ranging from .01 (white male comparison) to .03 (gender and Caucasian vs. African American analyses). Although the overall magnitude was not large, at least one item from each Reference and Focal comparison was rated moderate in DIF using the ETS rating system, and one item from the Caucasian vs. African American analyses was rated moderate on the basis of STD. Most of the items rated moderate in terms of DIF favored the Focal group.

GKQ0. Chi-square values were substantial for these analyses, especially for the gender comparison. Average values ranged from 1.03 for the white male analyses to 102.41 for gender. DD ranged from .00 to .16, for white males and gender, respectively. Three items from the gender analyses were rated B and three were rated C using the ETS system; four were rated moderate using STD as an index. Although only one item from the Caucasian vs. African American analyses was rated B, four were rated moderate on the basis of STD. Finally, for the Caucasian vs. Hispanic analyses two were rated B on the basis of $Mh\chi^2$ and DD, and one item was rated moderate and one severe with respect to STD. The majority of the items rated moderate to severe in terms of DIF favored the Reference group, although this is not reflected in the mean DD values.

GKS2. Chi-square values were much lower for this General Knowledge test, ranging from .81 for the white male analyses to 4.48 for gender analyses. Mean DD ranged from -.08 (gender) to .07 (Caucasian vs. African American). One item from the Caucasian vs. Hispanic analyses was rated C in favor of the Reference group. Two items were rated moderate in DIF on the basis of the STD index in the gender analyses.

Discussion

The goal of this study was to assess the extent to which this set of computer-administrated cognitive information processing tests resulted in DIF for comparisons of males vs females, Caucasians vs. African Americans, and Caucasians vs. Hispanics, with a control comparison of white males. Results indicated that the majority of tests demonstrated little or no DIF. Where DIF occurred, the number of items favoring the

Reference group tended to equal the number favoring the Focal group.

Descriptive statistics revealed four tests with lower than expected internal consistency reliability estimates. Mean chi-square values for three of the tests (INV1, INS2, and GKQ0) were substantially greater than for the remaining 11 tests, reflecting the tests' multidimensionality. Two of these tests, INV1 and GKQ0, demonstrated higher levels of DIF. An examination of the item content in these two tests suggested that differential exposure was a factor in the probability of responding correctly. As an example, items from INV1 which concerned food had a greater probability of being answered correctly by females than males, while items concerning geographical knowledge had a greater probability of being answered correctly by males. Further investigation into these tests is warranted.

Two additional subtests resulted in levels of DIF higher than observed in the majority of the subtests: SLQ3 and SLV1. No clear cause has been identified. Because of the nature of the test, where participants are expected to improve over trials (items) and blocks, it is likely that the analyses conducted in this study are not entirely appropriate.

Several recommendations for future research are in order. An examination of responses for the items flagged as performing differently for various subgroups revealed instability at the lower ends of ability, especially for the substantially negatively skewed tests. Increasing the range of ability would permit a better comparison of group performance at all levels of ability. In addition, collecting a larger sample from the Hispanic population could reduce the magnitude of DIF observed; this subgroup was affected more than others by smaller numbers of individuals at the lower ends of ability, especially for the longer tests. Finally, comparing the average levels of DIF found in the APT test to those either in the ASVAB or the AFQT would permit an evaluation of the extent to which these cognitive information-processing tests result in relatively less DIF than more traditional tests of ability.

References

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In P. W. Holland & H. Wainer (eds.) Differential Item Functioning (pp. 3-24). Hillsdale, N. J.: Lawrence Erlbaum.

Camilli. G., & Shepard, L. A. (1994). Methods for Identifying Biased Test Items, Vol. 4. Thousand Oaks, CA: SAGE Publications, Inc.

Christal, R. E. (1991). Comparative validities of ASVAB and LAMP tests for logic gates learning. (AL-TP-1991-0031). Brooks Air Force Base, TX: Manpower and Personnel Division, AFHRL.

Cole, N. S. (1993). History and development of DIF. In P. W. Holland & H. Wainer (eds.) Differential Item Functioning (pp. 25-30). Hillsdale, N. J.: Lawrence Erlbaum.

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (eds.) Differential Item Functioning (pp. 35-66). Hillsdale, N. J.: Lawrence Erlbaum.

Holland, P.W., & Thayer, D.T. (1988). Differential item functioning and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), Test Validity (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum.

Kyllonen (1994). CAM: A theorietical framwork for cognitive abilities measurements. In D. K. Detterman (ed.) Current topics in human intelligence (Vol. 4) (pp. 307-359). Norwood, NJ: Ablex Publishing Co.

Kyllonen, P.C., Chrystal, R.E., Woltz, D.J., Shute, V.J., Tirre, W.C., & Chaiken, S. (1990). Cognitive Abilities Measukrement (CAM) Test Battery, Version 4.0. Unpublished test battery, Armstrong Laboratory, Brooks Air Force Base, TX.

Langley, P. (1983). Exploring the space of cognitive architectures. Behavior Research Methods and Instruments, 15, 289-299.

Linn, R. L., Hastings, C. N., hu, P. G., & Ryan, K. E. (1988). Armed Services Vocational Aptitude Battery: Differential Item Functioning on the High School Form. (AFHRL-TR-87-45). Brooks Air Force Base, TX: Manpower and Personnel Division, AFHRL.

Rogers, H. J., & Hambleton, R. K. (1993). MH: A Fortran 77 program to compute the Mantel-Haenszel statistic for detecting differential item functioning. To appear in Educational and Psychological Measurement.

Welsh, J. R., Androlewicz, T. R., & Curan, L. T. (1990). Armed Services Vocational Aptitude Battery (ASVAB): Analyses of Differential Item Functioning on Forms 15, 16, and 17. (AFHRL-TP-90-62). Brooks Air Force Base, TX: Manpower and Personnel Division, AFHRL.

SKILL RETENTION AND DECAY: METHODOLOGICAL AND CONCEPTUAL ISSUES

Pamela L. Stanush
Graduate Student
Department of Psychology

Texas A&M University
College Station, TX 77843-4235

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

and

Armstrong Laboratory

August 1995

Skill retention and decay:
Methodological and conceptual issues

Pamela Stanush
Graduate Student
Department of Psychology
Texas A&M University

The overall objective of the following paper is to examine conceptual and methodological issues that should be addressed in research investigating skill decay and retention. Previous research in this area has suffered from weaknesses and inconsistencies that make drawing literature difficult. The following issues will be discussed: The importance of skill decay and long-term retention from the perspective of the military and its current training programs and protocols; and the methodological and conceptual problems with the variables and research designs that have been used in past and current skill retention/decay research. Suggestions for future research and a proposal for two follow-up studies that attempt to investigate the relationship between individual differences and long-term retention of skill decay are included.

SKILL RETENTION AND DECAY: METHODOLOGICAL AND CONCEPTUAL ISSUES

Pamela L. Stanush

INTRODUCTION

Millions of dollars each year are spent by private industry and the military to train personnel in the knowledge and skills needed for successful job performance (Lane, 1986). Unfortunately, the acquisition of skills and knowledge does not guarantee retention of that skill over time, especially in those situations in which skill and knowledge is not used or practiced. Instead, the robust finding that over time skill performance gradually falls below that achieved immediately following skill acquisition suggests that, without exercise, skill and knowledge progressively deteriorate or are gradually forgotten. Identifying the variables and conditions that influence or predict skill loss over time has many practical and financial implications

Research investigating the phenomena of skill retention, loss, and re-acquisition, however, although including a wide variety of variables, has suffered from a lack of standardization and consistent operationalization. This has resulted in a difficulty in identifying trends or drawing solid conclusions from the literature. Improvements and refinements in the methodology used to investigate the domain of skill decay and retention might have big payoffs if researchers interested in this domain take such suggestions into account when designing future research.

Skill loss during periods of non-use is particularly problematic for the military where it often occurs that military personnel are initially trained and then expected after a period of non-use to be able to use previously acquired knowledge or skills. Reserve personnel, for example, may be provided formal training only a couple of times a year with the expectation (or hope) that if these personnel are called up for active duty, they will need only a limited amount of refresher training to reacquire any skill that has been lost. While total retention of acquired skill over long periods of non-use is unrealistic, any improvement in the amount and quality of skill attained over intervals of minimal or

complete non-use represents savings in time, money, and possibly lives in the military, where tens of thousands of men and women undergo training every year.

There are several shortcomings and weaknesses present in the current skill decay literature. For example, the lack of standardization and/or consistent operationalization of key variables related to skill decay and the inconsistencies in methodology have made drawing conclusions about the phenomena of skill acquisition and retention difficult. These problems include a lack of adequate attention to measurement and conceptualization of skill acquisition, inconsistent operationalizations of criterion measures, and the absence of a "theory-guiding framework" (Farr, 1987).

Although skill acquisition is an integral part of any investigation in skill retention (since a skill cannot be retained if it has never been acquired), there has been a lack of attention given to the method and measurement of skill acquisition and its significance in determining the amount of skill/knowledge retained. Schmidt and Bjork (1992), for example, criticize training and educational settings for treating learning (i.e., skill acquisition) and retention as two separate phenomena which have been studied independently using different methods by different scientists and in different laboratories. Schmidt and Bjork argue that the two are really inseparable and need to be considered together when conducting studies on skill decay/retention.

Another weakness of the majority of skill retention/decay studies is the lack of attention given to pre-testing individuals for previous experience or skill. Since the majority of skill decay studies do not conduct a pre-test, it is difficult to ascertain whether a particular individual's learning curve begins at a knowledge or skill level of zero or whether it is actually higher.

The various criteria used to measure skill acquisition and the lack of standardization of these criteria is a methodological weakness in the skill retention/decay literature. Speed and accuracy are commonly used as criteria for measuring performance, although they have been criticized as being inadequate criterion measures to measure skill acquisition (Regian & Schneider, 1990). Accuracy, while an important measure, tends to asymptote quickly, and results in a unreliable measure of skill acquisition since learning continues beyond error-less performance. Speed allows for a more continuous measure of

change, but leaves the researcher with the problem of identifying and setting an appropriate criterion point at which to terminate skill acquisition. Regian and Schneider (1990) believe that even speed and accuracy together may not be sufficient to determine mastery and a third component, resource load, should be added, resulting in a definition of criterion performance as the point when "component speed, accuracy, and resource load have stabilized at a predetermined level."

Another issue that has been overlooked in the skill retention/decay literature is the characteristics associated with termination of skill acquisition and the learning curve. Since the learning rate is curvilinear, where an individual is on the learning curve when skill acquisition is terminated may have implications for the amount of skill retained over time. Jones (1985), for example, has argued that retention is adversely affected by terminating training prior to the period that performance has become relatively stable. To date, the majority of studies in the skill retention/decay domain have not systematically investigated issues related to termination of skill acquisition. Instead of empirically-derived or conceptually-based termination criteria, the skill retention literature have tended to use fairly arbitrary methods to determine when skill acquisition should be determined, and operationalizations of termination points and terminology have differed across studies. Many research paradigms investigating skill retention/decay have attempted to define proficiency and mastery in skill acquisition in order to investigate the phenomena of overlearning which is defined as additional training beyond that required for initial proficiency. Proficiency, for example, is often defined as one errorless trial (Hagman, 1980a; Hagman, 1980b) although one errorless trial has been labeled differently in other studies (e.g., "minimal mastery," Farr, 1987; and "mastery," Hall, Ford, Whitten, & Plyant, 1983). The term "mastery" has been operationalized in several ways. Mastery has been used to refer to one errorless trial (Hall et al., 1983), to two error-free trials (Schendel & Hagman, 1980), and to three error-free trials (Goldberg, Drillings, & Dressel, 1981).

By establishing a criterion such as one or three errorless trials that individuals much achieve before beginning the retention interval, researchers have attempted to standardize the amount of skill acquisition across individuals. There are several problems,

however, with this methodology. First, performance can be measured on several dimensions, of which accuracy is only one criterion. Secondly, the often arbitrary criterion level used to define proficiency and mastery by many skill retention/decay studies has implications for investigations about "overlearning." If skill acquisition and learning continue beyond that of one error-free performance, than investigations of overlearning may be confounded with initial skill acquisition leading to a difficult interpretation of where skill acquisition is considered complete and overlearning begins.

One exception to the previously mentioned studies is an empirical investigation by Jones' (1989) who investigated individual learning rates, overlearning, and skill retention. Jones statistically controlled the end-of-practice level (operationalized as the point at which performance on video games reached asymptote for each subject) in order to investigate skill retention. The study found that skill was retained for those individuals who overpracticed or who were more slowly improving late at practice. Although there were some weaknesses in his study, Jones' investigation is a step in the right direction. By taking into account when performance asymptotes and using this criterion as mastery, a more methodologically sound investigation of variables such as individual differences in rates of learning and overlearning resulted.

Jones (1989) found that individuals learn at the different rates which is a finding consistently supported by other investigations in the learning literature. Many factors affect the rate of learning including individual differences, task characteristic, instructional strategies, etc. Most of these variables can be recognized as being the same as those previously identified as moderators of the phenomena of skill decay. Individual differences appear to influence skill acquisition not only in quantitative terms (a greater amount) but also qualitatively (a "stronger" memory trace).

Many of the previously mentioned issues indicate that there are important individual differences related to skill acquisition that need to be taken into account when studying the effect of individual differences on skill retention. Several individual difference variables have been investigated in the skill decay literature as having the potential to influence long-term retention. These variables include age, previous experience, and motivation although the majority of individual differences studies have

focused on ability, or more broadly defined, general intelligence. Although several studies have found that higher ability individuals retain more skill and/or knowledge over time, most authors have concluded that increased retention by higher ability individuals can be explained by the increased amount of skill acquired before the retention interval began. Research investigating the influence of ability on retention of motor tasks have used military tasks (Goldberg et al., 1981; Grimsley, 1969; Vineberg, 1975) and using non-military-related tasks (Carron, 1971; Carron & Marteniuk, 1971; Purdy & Lockhart, 1962) have found similar results and drawn similar conclusions - individuals with higher initial ability retain a higher degree of skill over time because they either learn faster and acquire more skill in a specified amount of time compared to lower-ability individuals.

More recently, it has been recognized that since individuals of differing ability use different learning strategies when acquiring skills, it seems likely that higher ability individuals not only differ in the amount of skill that is attained but also may obtain a higher qualitative difference in skill that may enhance long-term retention. Farr (1987), for example, suggests that differential decay rates should occur between individuals with differing abilities because higher ability individuals may use more effective learning strategies to acquire skill. The conclusion by several authors that individuals of various abilities lose skill at the same rate seems counter-intuitive and could be a result of methodological problems already discussed that pervade the current studies investigating such issues.

The problem of defining mastery and the arbitrary nature by which it is defined in the skill decay/literature must be remedied in order to study individual differences. Jones (1985) proposes that skill acquisition should not be terminated until the learning curve begins to level off indicating that the skill is truly "mastered." Based on findings from an empirical investigation, Jones suggested that a more stable and accurate prediction of long-term retention can be attained by examining the shape of an individual's learning curve along with the trainee's approximation to some arbitrary criterion. Thus, one way to standardize skill acquisition might be train individuals until performance asymptotes for a particular task. The point at which performance asymptotes on average across individuals could represent a more standardized criterion for when skill acquisition would

end. Individuals would end training at this point and would later be tested for retention of skill after a predetermined length of time. The length of time to train to the average asymptote will vary by individuals but the quantity of skill acquisition should be relatively constant across individuals. An unconfounded relationship between individual differences such as cognitive ability, age, etc., and long-term acquisition could then be assessed.

PROPOSAL

Two studies are proposed that will investigate the relationship between specified individual differences and skill retention over a predetermined retention interval. First, a pilot study will be conducted to determine the average performance level at which a leveling of the learning curve occurs across individuals (i.e., when performance asymptotes). The second study will use the empirically determined performance level as a standardized level of skill acquisition, and individuals will be trained until they reach the predetermined criterion. Several individual difference constructs will be measured that are hypothesized to affect the quality of skill acquisition which in turn is hypothesized to effect the long term retention and reacquisition of the particular skill. After a predetermined retention interval, individuals will be tested for skill retention, and then will retrain to previous performance levels.

Both studies will utilize a complex psychomotor task called Space Fortress (Gopher, 1992; Shebilske, Regian, Arthur, & Jordan, 1992) which was designed as an experimental task that could be used to investigate the study of complex skill and its acquisition. Space Fortress has been used several times in investigations of the relationship between individual differences and skill acquisition. Foss, Famiani, Mane, & Donchin (1989) investigated the performance of 40 male subjects who were given no training information other than the standard game instructions. Subjects were initially given an aiming task which was developed as a screening tool and has been found to correlate positively with the subject's subsequent performance on Space Fortress. The subjects then were allowed to play the game for a period of ten sessions, each lasting one hour; sessions were run on consecutive days. Results indicated that performance of the game improved steadily throughout the training period and that performance did not appear to reach asymptote on any of the dependent variables. The aiming screening task was a good

predictor of the subjects' final overall performance on the game, and differences between subjects that manifested themselves first on the aiming task were evident at the beginning of training and clearly maintained over time. Individuals not only differed on final performance but also on learning rate and in style of play or strategy.

Rabbit, Banerji, and Szymanski (1989) investigated whether specified individual differences such as conventional pencil and paper IQ tests, age, and previous experience of video-games could predict performance on Space Fortress. Three levels of performance were assessed: initial performance, rate of learning, and maximum performance during five days of practice. Rabbitt, Banerki, & Szymanski (1989) found that previous experience predicted initial performance on Space Fortress but not rate of learning nor the maximum score achieved after the five days practice. Age predicted maximum performance but not the initial rate of improvement. Intelligence did not predict initial performance after the effects of age and previous experience were partialled out, but predict rate of subsequent learning. The results of the present study indicate that IQ tests can predict performance on a complex psychomotor task; surprisingly, little research has investigated the predictive ability of intelligence on the performance of psychomotor tasks. This research suggests that more intelligent individuals appear to be able to "master complex systems of rules, to attend selectively to the critical portions of complex scenarios, to make rapid and correct predictions of immanent events and to prioritize and update information in working memory" (pg. 255).

The previous research supports the hypothesis that at least some individual differences appear to not only affect the amount of skill acquisition acquired but also the quality as well. If the quality of skill acquisition is enhanced, then long term retention may also be enhanced independently of the quantity of skill acquisition acquired. One experiment to date has used the Space Fortress paradigm to investigate the influence of individual differences on skill retention over an eight-week retention interval and on re-acquisition of the skill (Arthur, 1994). This experiment investigated the effect of both dyadic and individual training protocols in the acquisition of skill; subjects were trained for one hour a day for two weeks. Individual differences included in the study included cognitive abilities, visual selection attention, spatial working memory, spatial processing

speed, social desirability and self-monitoring, interaction anxiousness, and personality measures. Although the results of the study are not completed, this study represents a first step in the investigation of the relationship between individual differences, most importantly cognitive abilities, and retention and reacquisition of a complex skill.

The two proposed studies will attempt to clarify the relationship between individual differences and skill retention by using a slightly different paradigm than previous studies using Space Fortress. Additionally, the proposed studies will be used to determine empirically a criterion that can be used as a standardized operationalization of mastery that will serve as the terminating point of skill acquisition. These two studies will complement the findings of the Arthur (1994) study and should provide some information that will allow for enhanced interpretation of any subsequent research utilizing the Space Fortress paradigm.

As mentioned earlier, skill retention studies have chosen rather arbitrary criteria for when skill acquisition should be terminated. For studies investigating individual differences of skill retention, this lack of standardization becomes particularly problematic since there are individuals differ in learning rate and time-to-learn. Thus, previous studies have confounded individual differences in acquisition with individual differences in retention and an uncontaminated relationship between individual differences and skill retention has not been attainable.

If individuals are all trained to the same criterion level of performance, such as mastery, than the relationship between individual differences and skill retention could be assessed. The problem is that mastery has been defined in a variety of ways, most of which are arbitrary and do not truly represent "mastery" as defined as perfect performance. Mastery can be defined as no errorless trials for dependent variables that have a definite endpoint or can be defined as mastery if the variable reaches an asymptote at the highest level of performance (score at the end of specified interval) or at the quickest time (errorless trial in the least amount of time). The definition of mastery chosen thus determines the training schedule. Training on the task can be manipulated in several ways. The majority of experiments using Space Fortress have involved training subjects for a fixed amount of time and then determining their performance level.

Although there is no avoiding confounding the time-on-task with the level reached by the subject at any given time, a more theoretical based point at which to end skill acquisition is when performance asymptotes. Investigations in skill acquisition have equated the leveling off of the learning curve, (i.e., when performance asymptotes), as occurring when consistent (or consistent parts of) tasks have become automatized, or require little cognitive attention. If the average level of asymptotic performance is determined for a particular task than this represents "mastery" such that learning is not increasing over time. This performance level than can represent a standardized operationalization of mastery and can be used as a termination point for skill acquisition across individuals.

Study 1: The pilot study will have the following objectives:

1) To provide an empirically derived level of skill acquisition that will be used as a standard criterion for determining when to consider skill acquisition complete.

2) To assess whether specified individual differences can predict the time to reach criterion and the final level of asymptotic performance.

After the level of asymptotic performance is determined, subsequent studies investigating individual differences in either skill acquisition or skill retention can use it as a standardized criterion of mastery. Individual differences such as cognitive ability and rate of learning can be assessed in terms of their relationship with skill acquisition, skill retention, time to retrain, etc.

METHOD

Subjects.

The sample will consist of 20 individuals recruited from a large southwestern university. All subjects will be male and right-handed. Subjects will be recruited by posted notices around the university and through advertisements in the school newspaper. Subjects will be told that they will paid $5 per hour for participation in the study. Subjects will practice one hour a day for up until their scores reach an asymptote. Asymptotic performance will be calculated as an improvement of less than 5% for two consequent days. Bonuses will be given to those subjects who reach asymptotic performance in the fewest number of days.

Predictors:

Computer-Administered Visual Attention Test (CA-VAT). The CAT-VAT (Arthur, 1991; Arthur, Strong, Jordan, Williamson, Shebilske, & Regian, in press; Arthur, Strong, & Williamson, 1994) is a test of visual attention which is computer administered and scored. It is a visual counterpart to the Auditory Selective Attention Test (ASAT; Gopher & Kahneman, 1971; Mihal & Barrett, 1976).

Computer-Attitude Scale (CATT). The CATT (Dambrot, Watkins-Malek, Silling, Marshall, & Garver, 1985) will used to assess the respondent's attitude toward computers. This instrument is a 20-item measure that is computer-administered.

Declarative Knowledge Test. This test measures the subjects' knowledge about the Space Fortress instructions and procedural rules. It is a paper-and-pencil measure with 30 multiple-choice items.

Figure Matrices Test (g; Kyllonen, Christal, Woltz, Shute, Tirre, & Chaiken, 1990). This measure is a computer-administered cognitive ability test that is comparable to the paper-and-pencil Raven Advanced Progressive Matrices test (Raven, Court, & Raven, 1985). It was developed at Brooks AFB as part of the Cognitive Abilities Measurement (CAM) test battery (version 4.0). The instrument consists of nine, homogenous, progressively difficult items.

Spatial Working Memory Test (SWMT; Kyllonen et al., 1990). This test is a 24-item test of spatial (working memory) that is computer administered. This instrument is also part of the CAM .

Spatial Processing Speed Test (SPST; Kyllonen et al., 1990). This test is a 12-item test of (spatial processing speed). It is computer-administered and is also part of the CAM.

Video Game Experience Questionnaire. This measure is designed to assess the subjects' video game experience prior to participation in the Space Fortress experiment and during the retention interval.

Confidence/Alertness Questionnaire (CAQ). This measure consists of 12 self-report items that are computer-administered. The instrument will be used to assess the

subjects' fatigue, motivation, and confidence in the performance on the Space Fortress Task.

Criteria:

Space Fortress. This paradigm was designed as an experimental task that could be used to investigate the study of complex skill and its acquisition. The game is a video-game like simulator in which the object is to shoot missiles at and destroy a space fortress using a joy-stick. In the process of destroying the fortress, the subject must overcome several obstacles. The subject's performance is evaluated in the form of a game score which is score of overall proficiency. It represents the sum of the points received for damage to hostile elements minus the penalties for damage to the spaceship. The subject is require to use a variety of skills including perceptual, cognitive, and motor skills. Specific knowledge of the rules and game strategy are also required.

The task was designed to be representative of real-world tasks, to incorporate various dimensions of difficulty that are relevant to research on skill acquisition, and to be stimulating enough to keep participants interested during extended practice. The task involves a multitude of skills including motor, visual monitoring and scanning, and memory requirements. Fifty parameters can be manipulated to define the shape and form of the game. Examples of parameters include the speed of the hostile elements and the time allowed before the fortress begins to fire at the subject's ship. A measure of overall proficiency is the final game score which is the sum of the points received for damage to hostile elements minus the penalties for damage to the spaceship. This final game score is used to evaluate performance of each subject.

The equipment required to use the task is an IBM AT compatible with an 80386 processor and a math co-processor, a VGA monitor, a joystick, and a three-button mouse. For additional information on the game refer to Gopher (1992) and Shebilske et al. (1992).

PROCEDURE

Screening.

Subjects who are recruited via advertisements and posted notices will be initially screened to ensure appropriateness for the study. Subjects will be informed that the study

consists of performing a video game-like task and completing several individual differences measures. Informed consent forms and contract for pay forms will be signed first before the screening session begins. Subjects will be screened using an aiming task that has been found to predict future performance both in terms of rate of learning and final performance (Mane, Coles, Daris, Strayer, & Donchin, 1984). A final score of at least 780 must be attained for the aiming task in order for a subject to continue in the study. The aiming task provides a task to screen out those subjects, who based on past research (Donchin, 1989) will find the task so difficult that they will have to be removed from the experiment. Thus, those individuals who do not score a minimum of 780 on the aiming task will not be allowed to continue. Additionally, any subject who plays more than 20 hours of videogames per week will be dismissed. The screening criteria is used to reduce error variance and to make subjects more representative of those in operational training centers.

Training and Test Sessions.

Training on the task can be manipulated in several ways. The majority of experiments using Space Fortress involve training subjects for a fixed amount of time and then determining their performance level. Another less frequently used approach to training on Space Fortress is to define the performance level that must be reached to measure the time to reach that level. As usual with skill acquisition tasks, there is no avoiding the time-on-task with the level reached by the subject at any given time. The approach to be used by the pilot study is a variation of the second less frequently used training paradigm. Subjects will be trained not to a predetermined performance level but instead will be allowed to practice until their performance asymptotes, that is, until the learning curve levels off. This level will be used to determine individually the point at which the task has been "mastered," as operationally defined in the pilot study. An average of the performance levels attained by the subjects in the pilot study will be used as a termination point for skill acquisition in the second study.

The second study will be investigate similar individual difference variables as Arthur (1994), however, the practice protocol will differ from Arthur in order to assess an unconfounded relationship between individual differences and skill retention and re-

acquisition. While Arthur (1995) standardized the amount of training received across subjects at approximately one hour of practice per day for two weeks, the proposed study will used the empirically determined termination point from the pilot study as the point at which to end skill acquisition for all subjects. Subjects will train for as long as it is needed to reach the predetermined performance level.

Study 2: The second study will have the following objectives:

1) To use the previously determined standardized level of skill acquisition to investigate the relationship between specified individual differences and skill retention over a predetermined retention interval
2) To investigate whether individual difference constructs influence the amount of training needed to reacquired skill originally attained.

METHOD

Subjects:

The sample will consist of 20 individuals recruited from a large southwestern university. All subjects will be male and right-handed. Subjects will be recruited by posted notices around the university and through advertisements in the school newspaper. Subjects will be told that they will paid $5 per hour for participation in the study. Subjects will practice one hour a day until their scores reach the predetermined performance level. Bonuses will be given to those subjects who reach the predetermined performance in the fewest number of days.

Predictors:

The same measures will be used as in the pilot study: Computer-Administered Visual Attention Test (CA-VAT), Computer-Attitude Scale (CATT), Declarative Knowledge Test, Figure Matrices Test, Spatial Working Memory Test (SWMT), Spatial Processing Speed Test (SPST), Video Game Experience Questionnaire, and the Confidence/Alertness Questionnaire (CAQ).

Criteria:

Space Fortress. Space Fortress will be used as the training task.

Asteroids. Asteroids (Logg, 1993) will serve as a near transfer task. It is a game that required that subjects control a spaceship, shoot asteroids, and maneuver through an asteroid belt. The game requires an IBM PC/AT computer with a standard keyboard.

Tempest. Tempest (Theurer, 1993) is a game that requires subjects to shoot and destroy "aliens" in a ship they control called "Blaster." This game requires an IBM PC/AT computer with a standard keyboard and mouse.

Space Fortress Keyboard Version. The keyboard version of Space Fortress is identical to the original version except that the joystick operations are replaced with the keyboard. This modification of the normal Space Fortress allows the keyboard version to be a transfer task to the original game.

PROCEDURE

Screening. Subjects who are recruited and screened in the same manner as that described in the pilot study. The study requirements, however, will be more extensive than that of the pilot study since these subjects will be trained until they reach the criterion performance level, then will return after eight weeks retention interval and retrained to criterion.

Training and Test Sessions. Subjects will be paid according to the number of hours of practice that are required to reach the predetermined performance level. After a subject reaches the standardized level, a eight-week retention level will follow. Subjects will also be competing for bonuses that will be awarded to those subjects who reach the criterion performance in the fewest number of days. The use of bonuses will hopefully discourage subjects from deliberately trying to not reach the criterion level of performance. Those subjects who do not reach the criterion level in a prespecified amount of time will end practice at that performance level. The data collection procedures will be similar to those in Arthur (1995).

After an eight-week retention interval, subjects will return to be tested for retention. Subjects will be tested on the transfer tasks, Asteroids, Tempest, and the Space Fortress keyboard version, and will be trained on Space Fortress in one hour sessions per day to test reacquisition of skill.

References

Arthur, W., Jr. (1991). Individual differences in the prediction and training of complex perceptual-motor skill tasks: The development and validation of the Computer-Administered Visual Attention Test. Tech. Report No. 8. College Station, TX: Texas A&M University, Psychology Department.

Arthur, W., Jr. (1994). Skill decay: A comparative assessment of training protocols and individual differences in the loss and re-acquisition of complex skills. Interim Report.

Arthur, W., Jr. (1995). Skill retention and decay: A meta-analysis. Presentation at the Xxth Annual APA Convention.

Arthur, W., Jr., Strong, M. H., Jordan, J. A., Williamson, J. E., Shebilske, W. L., & Regian. (in press). Visual attention: Individual differences in training and predicting complex task performance. Acta Psychologica.

Arthur, W., Jr., Strong, M. H., & Williamson, J. E. (1994). Validation of a visual attention test as a predictor of driving accident involvement. Journal of Occupational and Organizational Psychology, 67, 173-182.

Carron, A. V. (1971). Effect of ability upon retention of a balance skill after two years. Perceptual and Motor Skills, 33, 527-529.

Carron, A. V., & Marteniuk, R. C. (1970). Retention as a balance skill as a function of initial ability level. Research Quarterly, 41, 478-483.

Dambrot, F. H., Watkins-Malek, M. A., Silling, M. S., Marshall, R. S., & Garver, J. A. (1985). Correlates of sex differences in attitudes toward involvement with computers. Journal of Vocational Behavior, 27, 71-86.

Donchin, E. (1989). The learning strategies project: Introductory remarks. Acta Psychologica, 71, 1-15.

Farr, M. J. (1987). The long-term retention of knowledge and skills: A cognitive and instructional perspective. New York: Springer-Velag.

Foss, M. A., Fabiani, M., Mane, A. M., & Donchin, E. (1989). Unsupervised practice: The performance of the control group. Acta Psychologica, 71, 23-51.

Goldberg, S. L., Drillings, M., & Dressel, J. D. (1981). Mastery training: Effect on skill retention. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, Technical Report 513.

Gopher, D. (1992). The skill of attention control: Acquisition and execution of attention strategies. In D. Meyer & S. Korenblum (Eds.). Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience -- A silver jubilee. Cambridge, MA: MIT Press.

Gopher, D., & Kahneman, D. (1971). Individual differences in attention and the prediction of flight criteria. Perceptual and Motor Skill, 33, 135-1342.

Grimsley, D. L. (1969). Acquisition, retention, and retraining: Training category IV personnel with low fidelity devices (HumRRO Tech. Rep. 69-12). Alexandria, VA: Human Resources Research Organization.

Hagman (1980a). Effects of training schedule and equipment variety on retention and transfer of maintenance skill. Research Report 1309. Alexandria, VA: Army Research Institute.

Hagman (1980b). Effects of training task repetition in retention and transfer of maintenance skill. (Research Report No. ADA1201672XSP). Alexandria, VA: Army Research Institute for Behavioral and Social Sciences.

Hagman, J. D., & Rose, A. M. (1983). Retention of military tasks: A review. Human Factors, 25, 199-213.

Hagman, J. D., & Schendel, J. D. (1981). Effects of refresher training on job-task typewriting performance. JSAS Catalog of Selected Documents in Psychology, 11, 91. (Ms. No. 2391).

Hall, E. R., Ford, L. H., Whitten, T. C., & Plyant, L. R. (1983). Knowledge retention among graduates of basic electricity and electronics schools. (HumRRO Tech. Rep 149). Orlando, FL: Training Analysis and Evaluation Group, Department of the Navy.

Jones, M. B. (1985). Nonimposed overpractice and skill retention. (Contract No. MDA 903-83-K-0246.) Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Jones, M. B. (1989). Individual differences in skill retention. American Journal of Psychology, 102, 183-196.

Kyllonen, P. C., Christal, R. E., Woltz, D. J., Shute, V. J., Tirre, W. C., & Chaiken, S. (1990). Cognitive abilities measurement (CAM) test battery: Version 4.0 [Computer program]. Brooks Air Force Base, TX: Armstrong Laboratory.

Lane, N. E. (1986). Skill acquisition curves and military training. IDA Paper. Alexandria, VA: Institute for Defense Analysis.

Logg, E. (1993). Microsoft Arcade: Asteroids [Computer program]. Redmond, WA: Microsoft Corporation.

Mane, A. M., Coles, M. G., Karis, D., Strayer, D., & Donchin (1984). The design and use of subtasks in part training and their relationship to the whole task. Proceedings of the 28th annual conference on manual control.

Mihal, W., & Barrett, G. V. (1976). Individual differences in perceptual information processing and their relation to automobile accident involvement. Journal of Applied Psychology, 61, 229-233.

Purdy, B. J., & Lockhart, A. (1962). Retention and relearning of gross motor skills after long periods of no practice. Research Quarterly, 33, 265-272.

Rabbitt, P., Banerki, N., & Szymanski, A. (1989). Space Fortress as an IQ test? Predictions of learning and of practiced performance in a complex interactive video-game. Acta Psychologica, 71, 243-257.

Raven, J. C., Court, J. H., & Raven, J. (1985). A manual for Raven's Progressive Matrices and Vocabulary Scales. London: H. K. Lewis.

Regian, J. W., & Schneider, W,. (1990). Assessment procedures for predicting and optimizing skill acquisition after extensive practice. In N. Frederickson, R. Glaser, A. Lesgold, & M. G. Shafto (Eds.), Diagnostic monitoring of skill acquisition and knowledge acquisition (pg. 297-323). Hillsdale, NJ: Erlbaum.

Schmidt, R. A. , & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts in training. Psychological Science, 3, 207-217.

Shebilske, W. L., Regian, J. W., Arthur, W., Jr., & Jordan, J. A. (1992). A dyadic protocol for training complex skills. Human Factors, 34, 369-374.

Theurer, D. (1993). Microsoft Arcade: Tempest [Computer program]. Redmond, WA: Microsoft Corporation.

Vineberg, R. (1975). A study of the retention of skills and knowledge required in basic training. HumRRO-TR-75-10. Alexandria, VA: Human Resources Research Organization.

TEMPORAL PROCESSING: AN EXPLORATION OF TEST VALIDATION, COGNITIVE FACTORS, AND INDIVIDUAL DIFFERENCES

Mark J. Stimson
Doctoral Student
Department of Psychology

The University of Georgia
Athens, GA 30602

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1995

TEMPORAL PROCESSING: AN EXPLORATION OF TEST VALIDATION, COGNITIVE FACTORS, AND INDIVIDUAL DIFFERENCES

Mark J. Stimson
Doctoral Student
Department of Psychology
The University of Georgia

Abstract

The purpose of this study was threefold: to develop tests to measure temporal processing; confirm hypothesized latent factors of temporal processing; and investigate the relationship between temporal processing and working memory (i.e., individual differences). Human subjects were administered a series of different temporal processing tests that varied on several levels, as well as a battery of working memory tests. Several analyses were performed on the data. Inspection of the correlation matrix and a confirmatory analysis of the measurement model indicated construct validity for both temporal processing and working memory. Significance tests conducted on the temporal processing tests showed an increase in error with an increase in test duration. Finally, structural equation models confirmed hypothesized factors involved in temporal processing, and individual differences in working memory.

TEMPORAL PROCESSING: AN EXPLORATION OF TEST VALIDATION, COGNITIVE FACTORS, AND INDIVIDUAL DIFFERENCES

Mark J. Stimson

Introduction

The primary purpose of the Learning Abilities Measurement Program (LAMP) is to conduct research that examines the character of human performance. Previously, the LAMP research team had developed a series of computerized tests called the Cognitive Abilities Measurement (CAM) battery for measuring cognitive abilities. The principles used to successfully develop the CAM battery are now being used to accommodate psychomotor tasks to increase the predictability of performance in real-time multitasking situations. This new battery of tests is called PCAM. Furthermore, psychomotor testing is being reinstated into the Basic Attributes Test (BAT, Carretta, 1989) for selection of USAF pilots. Psychomotor testing is important because it measures complex real-time multitasking situations that encompass an array of human abilities. It is clear how psychomotor abilities are used in numerous military tasks such as operating piloting aircraft or construction equipment. The fundamental goal, therefore, is to develop a psychomotor test battery that will be useful in the selection, classification, and assessment of Air Force personnel.

To develop a battery of tests that accurately measures the whole spectrum of psychomotor skills, it helps to know which underlying factors are responsible for performance. Exploratory studies conducted in the past have found a taxonomy of latent factors. One early and influential investigation (Fleishman, 1964) found a set of 11 factors (e.g., response orientation, reaction time, and aiming) based on a series of perceptual-motor tasks. However, Fleishman's classic taxonomy may be outdated because of technological and statistical limitations. First, traditional psychomotor tests used static pencil-and-paper displays. Today, with the help of computers, more dynamic tests can be employed leading to a more accurate measurement of complex real-time tasks. For example, there is evidence that dynamic displays may require different skills than static displays (Hunt, Pellegrino, Frick, Farr, and Alderton,

1988). Second, Fleishman was not able to benefit from statistical analyses that would allow him to make an a priori hypothesis about the underlying psychomotor factors and compare it to empirical data (vis., confirmatory factor analyses). Therefore, a primary goal of the LAMP research group is to find the factors underlying psychomotor performance and then make tests that adequately measure those abilities.

The purpose of this present study is to propose and test for a hypothesized psychomotor ability called "temporal processing." Temporal processing is the ability to process temporal information (i.e., units of time). This includes "estimating" the amount of time passed, and can also include "extrapolating" that information and using it to judge the amount of time to pass. Several of Fleishman's (1964) psychomotor ability factors contained a temporal component, yet it was confounded with other processes (e.g., spatial). Investigating the possibility of an overall temporal factor, as well as building valid tests and looking for relations to their factors, can be extremely valuable for the LAMP project. For example, it is obvious that an operator of a vehicle must be able to estimate how much time has passed and/or how much time is needed until an event will happen.

The aim of the current study is to study temporal processing of a brief unit of time. There has been a vast amount of research on timing mechanisms, yet the majority has dealt with circadian rhythms (i.e., rhythms that work once a day, such as sleep). Little work has been done on briefer temporal periods called ultradian rhythms (Lovett Doust, Payne, and Podnieks, 1978). Of the studies that do investigate short temporal durations, many are confounded with spatial processing (e.g., Hunt et al., 1988; Peterken, Brown, and Bowman, 1991). That is, time is often tested using moving spatial objects. Furthermore, it is important to separate the processes of estimation and extrapolation (see, Irvine, Wright, Dennis, and Gould, 1991). Estimation can be thought of as tracking an event in time. Once the duration of the event has been processed, it can be brought forth to predict the duration of a future event. This later process is extrapolation

A battery of tests was developed to isolate the separate processes above. These processes are considered independent, except for their relationship through temporal processing. One test was similar to spatial-temporal tests used before by researchers such as Hunt et al. (1988) and Irvine et al. (1991). In these tests, subjects tracked an object moving across a computer screen.

At a certain point, the object would disappear at an invisible wall and the subject had to imagine that it was still moving in the same path and at a constant speed. The subject then had to indicate when it would reach a destination point by pressing a key. This test thus measures all 3 dimensions noted above. Spatial ability is measured because an image is tracked. Estimation is measured before the wall, and extrapolation is measured after the wall.

Another test was made that eliminated the spatial component of the traditional test. In this test, subjects would only see a digital clock in the center of the screen. The starting, stopping, and disappearing point were not places this time, but numbers or concepts. The clock would begin at 0, disappear at some point, and the subject had to judge when it would reach 100. Again, there was no spatial component to the task. Another test type isolated estimation by eliminating both the spatial component and extrapolation component. In this task, no tracking was required like before, but extrapolation was not necessary either because there was no wall. In this task a subject was simply instructed to count. Thus the subject just had to merely estimate the amount of time passed. In short, the 6 tests contained 1 test of spatial/extrapolation/estimation, 3 tests of extrapolation/estimation, and 2 tests of pure estimation. Each of these tests consisted of various sub-tests that manipulated duration and wall placement among other things. Furthermore, several tests of working memory were given to test for individual differences among the 2 domains. Working memory can be thought of as the most predictive of, and accounts for the most variance of, general intelligence tests (Kyllonen and Stephens, 1991). Therefore, it is assumed that working memory should be a good indicator of temporal processing.

Method

<u>The study</u>

The present study was part of an immense study that covered a period of 5 days (i.e., a normal week. Each subject participated in each test through the entire week. Table 1 lists how each test was given each day of the week. Briefly, a week of tests included the CAM battery (which included working memory), the PCAM battery, The Vocational Interest for Career Enhancement Instrument (VOICE), the Broadbent test, the Big Five Questionnaire and Self-

Description Survey (Christal, 1994), and finally the temporal processing tests. Mood surveys were also given at the beginning and end of each day. Again, the tests under investigation for the present study are temporal processing and working memory.

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
AM	Mood Survey PCAM Battery	Mood Survey CAM Battery	Mood Survey PCAM Battery VOICE Survey Broadbent Test	Mood Survey PCAM Battery	Mood Survey
PM	PCAM Battery (continued) Mood Survey	Drive Simulator Mood Survey	PCAM Battery (continued) Mood Survey	Big 5 Questionnaire Self-Description Temporal Processing Mood Survey	Big 5 Questionnaire Self-Description Drive Simulator Mood Survey

Table 1. Tests administered in larger study, ordered by time and day

Subjects

A total of 161 civilians was acquired through a temporary employment agency to serve as subjects. Civilians replaced actual Air Force personnel because of time restrictions placed on use of the military. To closely fit the population, all subjects were required to be between the ages of 18 and 26, and the male to female ratio was chosen to be 4 to 1. Subjects were paid 8 dollars per hour for participation.

Design

Each subject participated in each temporal processing test. Independent variables were wall (yes or no) and movement across the screen (i.e., spatial: yes or no). The dependent variable for the temporal processing tests was the time-accuracy error in clicking the mouse button to predict the time at which an event would reach its destination. A negative score represents underestimating the true time (i.e., responding early), and a positive score represents overestimating time (i.e., respond after the duration has passed). The higher the score absolute value of the score, the greater the error.

Materials

Computers. Subjects were run on IBM compatible computers at the Test Development Center at Lackland Air Force Base, San Antonio, TX. Each computer included a keyboard for response entry and a display monitor. All test materials, including presentation, feedback, and scoring programs, were written using a high-level cognitive authorial system called PLATS.

Working memory tests. As far back as 1938 (Thurstone), there has been evidence for separate dimensions within working memory. The CAM contains working memory tests that independently measure the verbal, quantitative, and spatial dimensions of working memory. Collectively, these sub-tests adequately measure the dimensions that represent working memory.

Temporal processing tests. A total of 6 different principal-tests consisted of various sub-tests. A total of 32 sub-tests was given in all. Table 2 shows a list of principal- and sub-tests. For the partial-line test, a line begins to move on the left side of the screen and continues to the right of the screen at 1 of 3 different rates. When the line has moved through either 25% or a half of the screen (depending on the sub-test), it disappears behind an invisible "wall." In the partial-clock test, a digital clock is shown at the center of the screen, begins at 0 and ascends to 100 at 1 of 3 speeds. When the digit reaches either 25 or 50 (again, depending on the sub-test) the digit disappears. In the partial-dot test, a colored circle of varying size flashes on the screen at 1 of 3 rates, and disappears at either 3 or 5 flashes. The partial-tone test is identical to the partial-dot test except that the event is auditory. In the whole-racer test, an animal (e.g., a horse) is shown at the left side of the screen with a digital number above it (either 0..5, 0..10, 0..15, or 0..20, depending on the level of condition). The whole-clock test is similar except that only a large digital clock is shown in the center of the screen.

Procedure

On Thursday afternoon, subjects began a series of temporal processing tests after completing the Self-Description Survey (Christal, 1994). Each subject completed a whole principal test before beginning another. The order of principal test was randomized. The order of sub-tests within each principal-test was also randomized so that the subject was unaware of the duration or wall placement would be from trial to trial. For each principal-test, subjects were instructed to click the mouse when they thought that the event had reached its destination. For

the partial-line test they were instructed to click the mouse when they thought that the line would reach a flag at the right side. For the partial-clock test, they were told to click when they thought it would reach 100. For the partial-dot and tone test, they were told to click when they thought it would reach a count of 10. Finally, for the whole-racer and whole-clock test, they were instructed to click when 5, 10, 15, or 20 sec had passed. Subjects were given feedback about whether they had succeeded. Each principal-test consisted of practice trials to reduce any effect of learning.

Principal Test	Wall	Spatial	Sub-Test: Duration of Event	Sub-Test: Place of Wall
Partial-Line	Yes	Yes	5.0 sec	25%
			5.0 sec	50%
			7.5 sec	25%
			7.5 sec	50%
			10 sec	25%
			10 sec	50%
Partial-Clock	Yes	No	5.0 sec	25%
			5.0 sec	50%
			7.5 sec	25%
			7.5 sec	50%
			10 sec	25%
			10 sec	50%
Partial-Dot	Yes	No	5.0 sec	33%
			5.0 sec	50%
			7.5 sec	33%
			7.5 sec	50%
			10 sec	33%
			10 sec	50%
Partial-Tone	Yes	No	5.0 sec	33%
			5.0 sec	50%
			7.5 sec	33%
			7.5 sec	50%
			10 sec	33%
			10 sec	50%
Whole-Racer	No	No	5 sec	----
			10 sec	----
			15 sec	----
			20 sec	----
Whole-Clock	No	No	5 sec	----
			10 sec	----
			15 sec	----
			20 sec	----

Table 2. Temporal processing tests; principal and sub

Results and Discussion

Data manipulation

Before any analyses were conducted, several subjects were removed from the data base because they had either failed to remain for the whole battery, or because they had performed the tests inadequately (e.g., constantly clicking the mouse button prematurely). The number of remaining subjects was 137. Next, error scores were transformed non-linearly (i.e., logged) to appropriately represent the importance of each error (see, Cohen and Cohen, 1983). For example, if the task requirement was to indicate when 10 sec had passed, an error of 9 sec early is more serious than 9 sec late. Logging each score helps balance the seriousness of each error. In fact, logging the scores made the overall distribution more normal. The kurtosis and the skewness of the logged data were 4.14 and 2.21, while they were 10.86 and 3.09 for the normal scores, respectively. Lastly, all outliers (i.e., 3 SD's from the mean of each subject) were removed to make the distribution less skewed (see, Ratcliff, 1993).

Preliminary analyses

First, one analysis of variance (ANOVA) was conducted to determine whether there were differences between colors in the partial-dot task, and another ANOVA was generated to test differences between pitches in the partial-tone test. No differences were found, as expected, so color and sub-tests were combined. Next, plots of accuracy and sequence were made for each sub-test to determine whether there were effects of learning. Each plot showed no tendency for errors to ascend or descend across trials. It appears that the practice trials raised performance level to its asymptotic level for subjects.

Construct validity

Table 3 below shows the correlation matrix for each temporal processing sub-test and working memory test. Validity is shown if correlations among sub-tests are higher within a principal-test, than sub-tests that do not share a principal-test. A quick inspection of the matrix reveals that this indeed is the case. Therefore, it is proposed that the orthogonal principal-tests of temporal processing are measuring what they were intending to measure. Another observation that can be made from the matrix is the relationship between the temporal processing and working memory tests. These correlations were expected to be negative because the dependent

variable of temporal processing is an error score while a high score on working memory indicates success.

	PL-D5 W25	PL-D5 W50	PL-D7.5 W25	PL-D7.5 W50	PL-D10 W25	PL-D10 W50	PC-D5 W25	PC-D5 W50	PC-D7.5 W25	PC-D7.5 W50	PC-D10 W25	PC-D10 W50
PL-D5W25	1.00											
PL-D5W50	0.70	1.00										
PL-D7.5W25	0.67	0.68	1.00									
PL-D7.5W50	0.63	0.59	0.71	1.00								
PL-D10W25	0.61	0.61	0.69	0.61	1.00							
PL-D10W50	0.57	0.51	0.63	0.68	0.63	1.00						
PC-D5W25	0.28	0.39	0.25	0.19	0.32	0.28	1.00					
PC-D5W50	0.32	0.30	0.16	0.19	0.27	0.30	0.73	1.00				
PC-D7.5W25	0.40	0.38	0.48	0.55	0.47	0.47	0.43	0.59	1.00			
PC-D7.5W50	0.19	0.28	0.30	0.32	0.24	0.29	0.49	0.53	0.54	1.00		
PC-D10W25	0.18	0.31	0.26	0.34	0.34	0.45	0.50	0.51	0.56	0.60	1.00	
PC-D10W50	0.18	0.25	0.31	0.34	0.27	0.38	0.50	0.67	0.64	0.65	0.67	1.00
PD-D5W33	0.38	0.47	0.47	0.39	0.43	0.37	0.49	0.50	0.54	0.50	0.41	0.54
PD-D5W50	0.40	0.50	0.51	0.37	0.52	0.38	0.54	0.54	0.55	0.52	0.34	0.52
PD-D7.5W33	0.22	0.40	0.41	0.33	0.44	0.31	0.51	0.54	0.50	0.64	0.50	0.63
PD-D7.5W50	0.37	0.43	0.47	0.39	0.43	0.32	0.46	0.51	0.54	0.63	0.41	0.54
PD-D10W33	0.22	0.36	0.37	0.28	0.38	0.37	0.48	0.52	0.46	0.55	0.46	0.53
PD-D10W55	0.31	0.46	0.47	0.34	0.47	0.38	0.52	0.52	0.56	0.60	0.41	0.53
PT-D5W33	0.44	0.56	0.53	0.45	0.50	0.40	0.44	0.40	0.49	0.48	0.43	0.43
PT-D5W50	0.40	0.49	0.38	0.33	0.40	0.35	0.50	0.48	0.44	0.43	0.40	0.40
PT-D7.5W33	0.35	0.48	0.52	0.45	0.41	0.38	0.36	0.32	0.43	0.47	0.37	0.36
PT-D7.5W50	0.26	0.47	0.48	0.31	0.42	0.31	0.38	0.34	0.35	0.46	0.36	0.36
PT-D10W33	0.38	0.50	0.49	0.42	0.43	0.36	0.42	0.44	0.52	0.56	0.46	0.48
PT-D10W55	0.43	0.58	0.50	0.48	0.42	0.40	0.34	0.33	0.43	0.42	0.39	0.39
WR-D5	0.24	0.25	0.26	0.22	0.26	0.37	0.41	0.50	0.40	0.40	0.49	0.41
WR-D10	0.26	0.32	0.32	0.26	0.27	0.37	0.42	0.45	0.39	0.37	0.49	0.43
WR-D15	0.20	0.24	0.27	0.20	0.29	0.35	0.43	0.46	0.32	0.39	0.46	0.41
WR-D20	0.18	0.24	0.25	0.22	0.27	0.38	0.50	0.52	0.32	0.50	0.51	0.44
WC-D5	0.23	0.23	0.27	0.23	0.33	0.32	0.34	0.23	0.14	0.24	0.35	0.24
WC-D10	0.33	0.31	0.26	0.31	0.37	0.36	0.39	0.28	0.12	0.26	0.34	0.22
WC-D15	0.29	0.29	0.34	0.34	0.34	0.41	0.40	0.33	0.21	0.35	0.38	0.32
WC-D20	0.32	0.36	0.35	0.31	0.42	0.43	0.40	0.28	0.16	0.26	0.36	0.23
WMV	-0.33	-0.28	-0.39	-0.34	-0.46	-0.39	-0.19	-0.23	-0.33	-0.19	-0.25	-0.27
WMQ	-0.45	-0.40	-0.42	-0.35	-0.52	-0.37	-0.28	-0.23	-0.27	-0.32	-0.28	-0.15
WMS	-0.21	-0.25	-0.29	-0.25	-0.36	-0.36	-0.28	-0.23	-0.24	-0.24	-0.32	-0.22

	PD-D5 W33	PD-D5 W50	PD-D7.5 W33	PD-D7.5 W50	PD-D10 W33	PD-D10 W50	PT-D5 W33	PT-D5 W50	PT-D7.5 W33	PT-D7.5 W50	PT-D10 W33	PT-D10 W50
PD-D5W33	1.00											
PD-D5W50	0.81	1.00										
PD-D7.5W33	0.81	0.79	1.00									
PD-D7.5W50	0.77	0.79	0.82	1.00								
PD-D10W33	0.67	0.67	0.77	0.71	1.00							
PD-D10W55	0.78	0.86	0.81	0.76	0.76	1.00						
PT-D5W33	0.63	0.61	0.61	0.65	0.56	0.63	1.00					
PT-D5W50	0.46	0.57	0.52	0.60	0.49	0.51	0.80	1.00				
PT-D7.5W33	0.54	0.57	0.57	0.60	0.51	0.59	0.82	0.67	1.00			
PT-D7.5W50	0.56	0.64	0.69	0.65	0.59	0.67	0.82	0.78	0.82	1.00		
PT-D10W33	0.62	0.66	0.66	0.68	0.60	0.71	0.88	0.74	0.81	0.85	1.00	
PT-D10W55	0.47	0.54	0.52	0.56	0.43	0.56	0.79	0.75	0.78	0.80	0.84	1.00
WR-D5	0.29	0.29	0.32	0.31	0.34	0.35	0.31	0.37	0.27	0.32	0.35	0.27
WR-D10	0.28	0.31	0.34	0.29	0.33	0.37	0.34	0.39	0.26	0.34	0.40	0.29
WR-D15	0.27	0.28	0.35	0.29	0.34	0.35	0.26	0.34	0.21	0.31	0.33	0.22
WR-D20	0.29	0.30	0.40	0.32	0.37	0.37	0.36	0.39	0.32	0.37	0.38	0.29
WC-D5	0.16	0.19	0.25	0.20	0.37	0.21	0.34	0.42	0.33	0.32	0.32	0.25
WC-D10	0.20	0.17	0.22	0.19	0.35	0.21	0.38	0.39	0.35	0.30	0.33	0.30
WC-D15	0.25	0.28	0.28	0.26	0.41	0.31	0.35	0.38	0.38	0.34	0.39	0.31
WC-D20	0.22	0.25	0.27	0.20	0.39	0.30	0.38	0.38	0.37	0.36	0.34	0.31
WMV	-0.26	-0.35	-0.25	-0.25	-0.23	-0.32	-0.28	-0.21	-0.32	-0.28	-0.30	-0.20
WMQ	-0.20	-0.29	-0.22	-0.25	-0.22	-0.30	-0.28	-0.24	-0.30	-0.22	-0.28	-0.26
WMS	-0.17	-0.23	-0.24	-0.19	-0.26	-0.27	-0.24	-0.19	-0.25	-0.18	-0.19	-0.10

	WR-D5	WR-D10	WR-D15	WR-D20	WC-D5	WC-D10	WC-D15	WC-D20	WMV	WMQ	WMS
WR-D5	1.00										
WR-D10	0.80	1.00									
WR-D15	0.83	0.83	1.00								
WR-D20	0.80	0.78	0.91	1.00							
WC-D5	0.34	0.36	0.40	0.39	1.00						
WC-D10	0.33	0.32	0.38	0.41	0.83	1.00					
WC-D15	0.42	0.38	0.49	0.50	0.81	0.88	1.00				
WC-D20	0.40	0.38	0.47	0.48	0.84	0.90	0.89	1.00			
WMV	-0.29	-0.29	-0.30	-0.30	-0.28	-0.27	-0.32	-0.34	1.00		
WMQ	-0.32	-0.32	-0.31	-0.36	-0.19	-0.27	-0.28	-0.34	0.47	1.00	
WMS	-0.48	-0.44	-0.43	-0.44	-0.29	-0.31	-0.35	-0.38	0.54	0.49	1.00

Table 3. Correlation matrix for temporal processing and working memory (PL=partial-line; PC=partial-clock; PD=partial-dot; PT=partial-tone; WR=whole-racer; WC=whole-clock; WM=working memory; D=duration; W=wall; WMS= spatial; WMQ= quantitative; and WMV= verbal)

Beyond inspecting the correlation matrix, a satisfactory measurement model was found using LISREL (Joreskog and Sorbom, 1993). Each path coefficient from the standardized solution for the relationship between observed variables (i.e., sub-tests) and the latent variables (i.e., principal-test) are high and significant (see Figure 1).

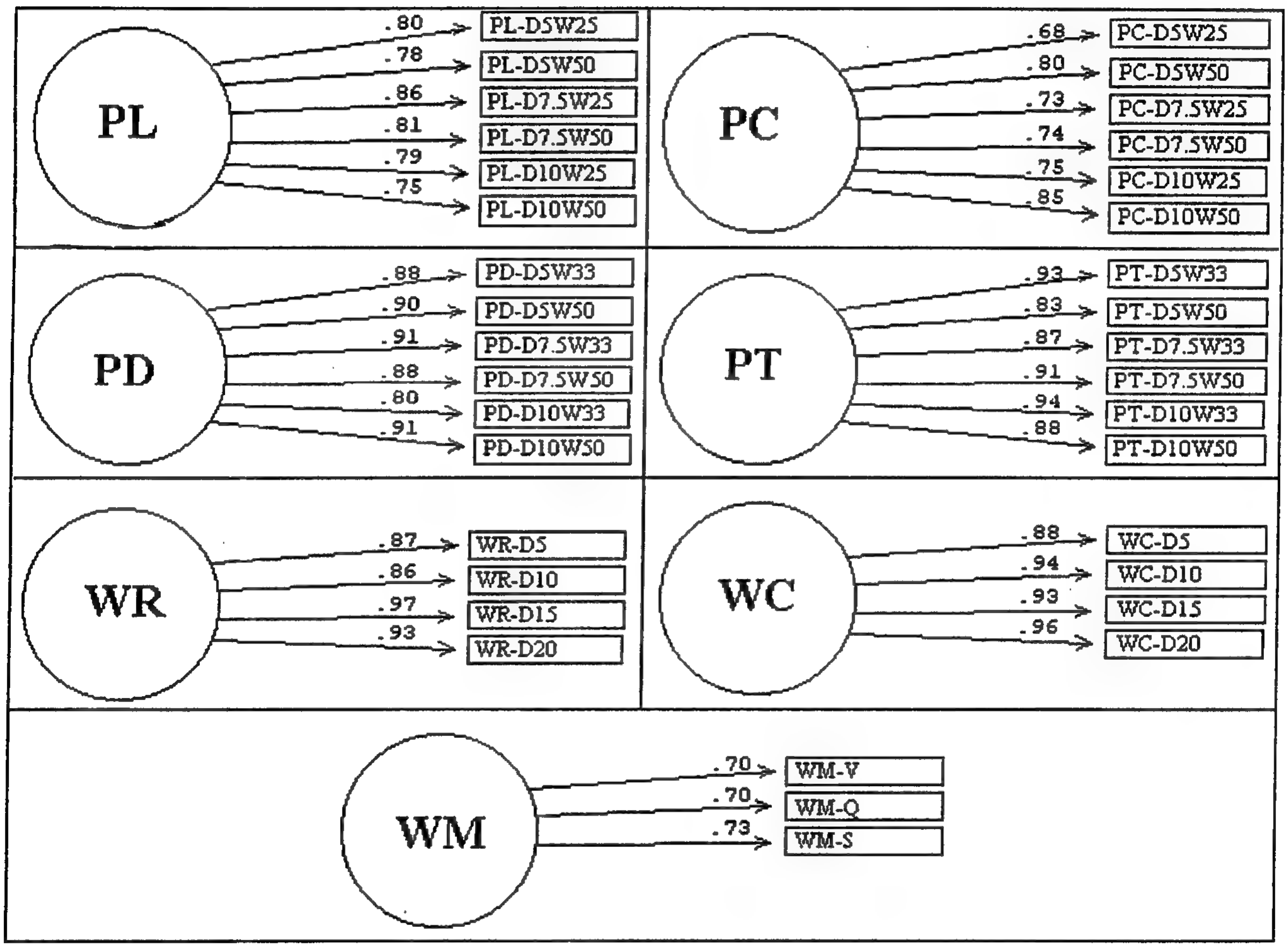


Figure 1. Measurement model for temporal processing and working memory (boxes represent observed variables; circles represent latent factors; values are standardized path coefficients; PL=partial-line; PC=partial-clock; PD=partial-dot; PT=partial-tone; WR=whole-racer; WC=whole-clock; WM=working memory; D=duration; W=wall; WMS=working memory spatial; WMQ=working memory quantitative; and WMV=working memory verbal)

Means and significance tests for temporal processing

The means and standard deviations are shown in Table 4. In each principal-test, the effect of duration was significant ($p<.01$; see Table 5). In each case, as the duration of the test increased, a person's accuracy in predicting time decreased. The means in Table 4 illustrates the rate of decline for each test. The finding that accuracy decreases with duration is not new, especially in spatial/temporal tasks when the extrapolation period is long (e.g., Dror, Kosslyn, and Waag, 1993; Gottsdanker, 1952; Jagacinski, Johnson, and Miller, 1983; Peterken, et al., 1991; Rosenbaum, 1955; Runeson, 1975). Peterken et al. argued that the increase in time-extrapolation, not time-estimation, is responsible for this trend. However, because extrapolation is often dependent on estimation, this conclusion may be tentative. The decline in performance was also found with the 2 principal-tests that did not require extrapolation. In a literature review of temporal processing, Allan (1979) pointed out that early studies by Vierordt in the 19th century showed that the estimation of short intervals are overestimated and long intervals are underestimated. This is indeed what was found in 3 of the 4 walls, or extrapolation, principal-tests.

The explanation of these results can be somewhat puzzling, depending on the chosen interpretation of the term "duration." One way to think about duration is that it is the amount of time to study the event. A plethora of cognitive research shows that the more an event is studied the better the retrieval process (e.g., Atkinson and Shiffrin, 1968). Therefore, it would be expected that an event would be studied more given more time, and that this would result in better accuracy as time increases. However, this is certainly not so here. Another way to think about duration is as time-units. Accordingly, duration is no longer a blessing to the processor, but a hindrance. As duration increases, so does the amount of time-units. Therefore, as time-units build up, the less room working memory has. This strain on working memory may then be responsible for the decline in accuracy as duration increases. These analyses do not necessarily support either argument because of the exploratory nature of the study and the fact that the hypotheses are post hoc. Nevertheless, the suppositions are helpful in grasping the situation better.

Temporal Processing Test	Mean	SD
Partial-Line Test	-.032	.103
Duration 5.0 sec	.007	.104
Wall 25%	-.002	.096
Wall 50 %	.015	.119
Duration 7.5 sec	-.033	.099
Wall 25%	-.025	.095
Wall 50 %	-.042	.103
Duration 10.0 sec	-.069	.093
Wall 25%	-.048	.081
Wall 50 %	-.090	.099

Temporal Processing Test	Mean	SD
Partial-Clock Test	-.016	.090
Duration 5.0 sec	.019	.088
Wall 25%	.027	.096
Wall 50 %	.012	.077
Duration 7.5 sec	-.019	.085
Wall 25%	-.028	.094
Wall 50 %	-.011	.074
Duration 10.0 sec	-.047	.085
Wall 25%	-.068	.093
Wall 50 %	-.026	.070

Temporal Processing Test	Mean	SD
Partial-Dot Test	-.012	.074
Duration 5.0 sec	.003	.078
Wall 33%	.001	.083
Wall 50 %	.005	.073
Duration 7.5 sec	-.009	.067
Wall 33%	-.013	.068
Wall 50 %	-.005	.066
Duration 10.0 sec	-.030	.074
Wall 33%	-.040	.076
Wall 50 %	-.019	.070

Temporal Processing Test	Mean	SD
Partial-Tone Test	-.013	.069
Duration 5.0 sec	-.012	.067
Wall 33%	-.015	.070
Wall 50 %	-.009	.063
Duration 7.5 sec	-.010	.071
Wall 33%	-.012	.073
Wall 50 %	-.007	.068
Duration 10.0 sec	-.016	.070
Wall 33%	-.022	.072
Wall 50 %	-.010	.067

Temporal Processing Test	Mean	SD
Whole-Racer Test	-.046	.111
Duration 5 sec	-.013	.105
Duration 10 sec	-.036	.099
Duration 15 sec	-.064	.113
Duration 20 sec	-.070	.119

Temporal Processing Test	Mean	SD
Whole-Clock Test	-.051	.115
Duration 5 sec	-.032	.111
Duration 10 sec	-.045	.112
Duration 15 sec	-.060	.112
Duration 20 sec	-.066	.121

Table 4. Means and SD's of each principal- and sub-temporal processing test.

Temporal Processing Test	df	F	p
Partial-Line Test	2, 2361	114.77	0.000
Partial-Clock Test	2, 2393	119.48	0.000
Partial-Dot Test	2, 4883	104.26	0.000
Partial-Tone Test	2, 4887	5.21	0.006
Whole-Racer Test	3, 1640	26.72	0.000
Whole-Clock Test	3, 1638	9.03	0.003

Table 5. ANOVA's for the duration condition (note, the 2nd degree of freedom value represents the number of items minus one)

Another independent variable that was found significant was "wall," (i.e., a shorter wall tended to produce fewer errors). In the present study however no attempt will be made to interpret these results because the amount of time before and after the wall, estimation and extrapolation respectively, is confounded.

Confirmatory factor analysis and individual differences testing

In Figure 2, the a priori causal model is shown. Each principal-test is part of a higher-order confirmatory factor analysis. That is, each principal-test variable is a 1st order factor and is related to one or more 2nd-order factor. The 2nd-order factors (i.e., spatial, estimation, and extrapolation) are theoretically motivated latent variables. As shown by the path coefficients, the partial-line test loaded high on the theoretical factor spatial ability. Each principal-test loaded significantly on a general time estimation factor. Finally, the 4 principal-tests that included a wall loaded high on time extrapolation. It should be noted that ultimately it would be better to have more than 1 first-order factor load on a higher-order factor, as is the case with partial-line on spatial ability. However, the design of the study does not permit this. Briefly, the variance of the principal latent test factor was fixed to 0.0. The path coefficient between the 1st-order factor and its 2nd-order factor, spatial ability, was then fixed to 1.0 so that the value of 2nd-order factor could represent the value of the 1st-order. This method is not confirmatory in the traditional sense, but the author feels that the spatial character of the partial-line sub-tests is so strong that it can be assumed that the latent factor does contain a good deal of a spatial component.

The other part of the structural equation model shows that the 2nd order latent variables are hypothesized to predict or cause the latent variable, working memory. This later character of the structural equation model is what constitutes the individual differences test. Working memory was found to load high on both the spatial and estimation factors. The high negative loadings suggest that as error in spatial- and estimation-temporal factors decreases, working memory performance increases. A test of working memory on extrapolation was also conducted, however the path-coefficient was very low and non-significant. The goodness of fit of the theoretical model to the empirical model is shown in Table 6. The root mean squared error of approximation value indicates an acceptable fit (see Brown and Cudeck, 1993). It should be noted that the fit is exceptionally good given the restrictions of the measurement model, and the

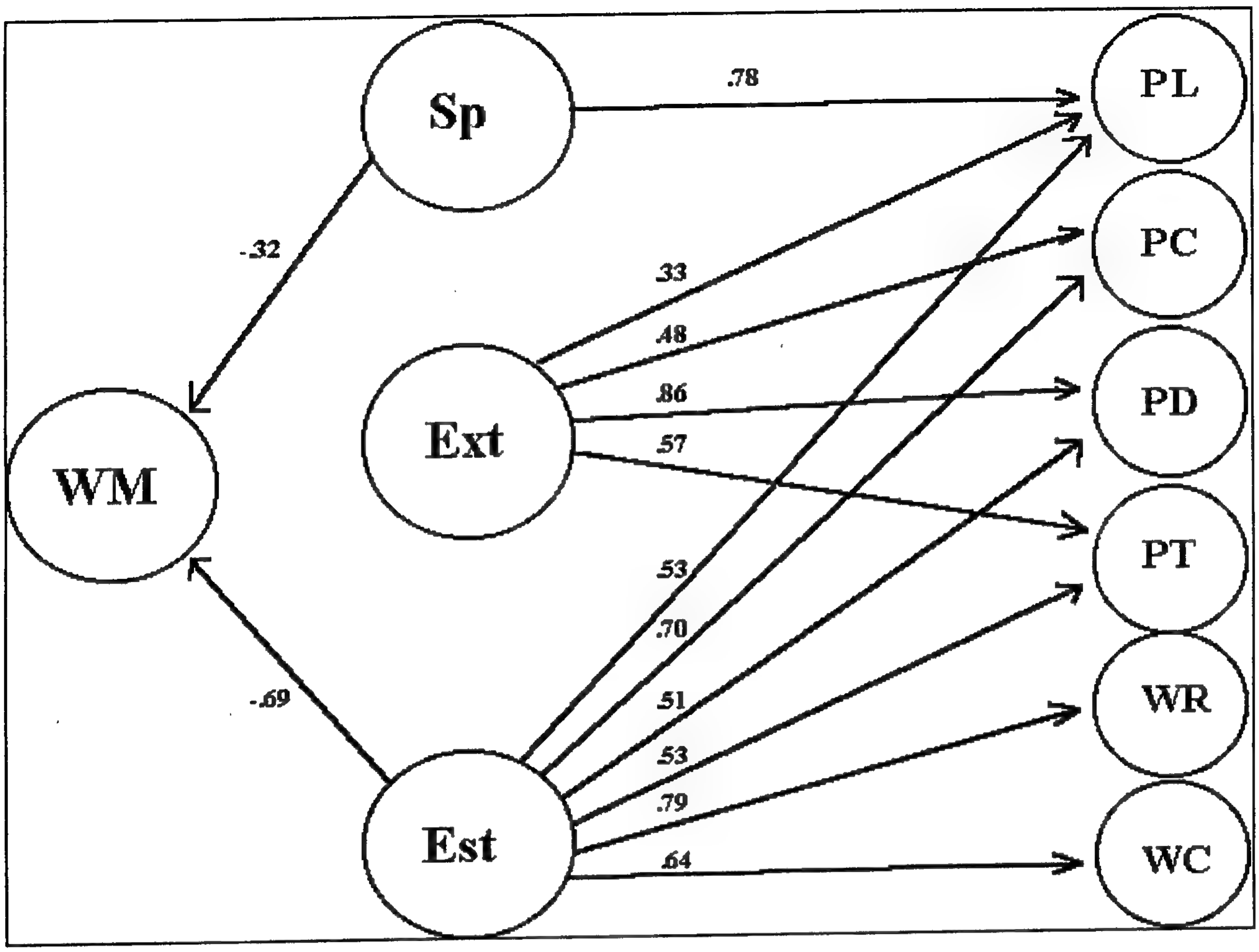


Figure 2. Structural equation model for temporal processing and working memory (circles represent latent factors; values are standardized path coefficients; PL=partial-line; PC=partial-clock; PD=partial-dot; PT=partial-tone; WR=whole-racer; WC=whole-clock; Sp=spatial; Ext=extrapolation; Est=estimation; and WM=working memory))

low sample size. That is, if all the latent factors of the measurement model were allowed to correlate, the best possible fit is obtained. If additional latent factors are attached, as well as additional paths, the fit can only decrease. Thus, if the theoretical model obtains a fit reasonably close to the correlated measurement model, then the prior model is strongly supported. Table 6 shows how close the fit is between the correlated measurement model and the theoretical model.

Theoretical Model		Mathmatical Model	
df	549	df	539
Chi-square	1115.78	Chi-square	1092.09
RMSEA	0.087	RMSEA	.087
GFI	0.69	GFI	.70
NFI	0.80	NFI	.80

Table 6. Goodness of fit indices for theoretical and mathematical model (df=degrees of freedom; RMSEA=root mean squared error; GFI=goodness of fit index; NFI=normed fit indix)

Conclusion

The present study was successful in three ways. First, the tests for temporal processing were found to satisfactorily measure what they were intended to measure. Second, the finding of a linear trend of error for duration supported prior research. Third, the structural equation model supported three theoretically motivated temporal processing latent factors, one working memory latent factor, and a finding of individual differences with respects to spatial/temporal processing and time-estimation with working memory. Thus, the fundamental properties of temporal processing have been explored and related to other processes used in testing Air Force recruits. Moreover, the present study suggests that there may be a general psychomotor factor of temporal processing. Future research should benefit from some of the design problems encountered. To strengthen the notion that there is a spatial component in temporal processing, more than one spatial principal-test should be implemented. Future experiments should also unconfound estimation and extrapolation. Finally, it might be interesting to use a wider range of duration levels to test Vierordt's law.

References

Allan, L. G. (1979). The perception of time. Perception and Psychophysics, 26, 340-354.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In W. K. Spence and J. T. Spence (Eds.), The psychology of learning and motivation: Advances in research and theory (Vol. 2, pp. 89-195). New York: Academic Press.

Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Editors): Testing Structural Equation Models. Sage Publications.

Carretta, T. R. (1989). USAF pilot selection and classification systems. Aviation, Space, Environmental Medicine, 60, 46-49

Christal, R. E (1994). Non-cognitive research involving systems of testing and learning (ASD) (Contract No. F33615-91-D-0010). Armstrong Laboratory, Brooks AFB, San Antonio, TX.

Cohen, J., Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Dror, I. E., Kosslyn, S. M., & Waag, W. L. (1993). Visual-spatial abilities of pilots. Journal of Applied Psychology, 78, 763-773.

Fleishman, E. A. (1964). The structure and measurement of physical fitness. Englewood Cliffs, NJ: Prentice-Hall.

Gottsdanker, R. M. (1952). The accuracy of predicted motion. Journal of Experimental Psychology, 43, 26-36.

Hunt, E., Pellegrino, J. W., Frick, R. W., Farr, S. A., & Alderton, D. (1988). The ability to reason about movement in the visual field. Intelligence, 12, 77-100.

Irvine, S. H., Wright, D. A., Dennis, I., & Gould, P. (1991). Time estimation, aiming and tracking-t: Measurement of individual differences in the performance of military tasks (HAL Technical Report: HAL3-91). Devon, UK. Human Assessment Laboratory.

Jagacinski, R. J., Johnson, W. W., & Miller, R. A. (1983). Quantifying the cognitive trajectories of extrapolated movements. Journal of Experimental Psychology: Human Perception and Performance, 9, 43-57.

Joreskog, K. G., Sorbom, D. (1984). LISREL-8.02: Analysis of linear structural relations by method of maximum likelihood. Chicago, IL: Scientific Software International, Inc.

Kyllonen, P. & Stephens, D. (1990). Cognitive abilities as determinants of success in acquiring logic skill. Learning and Individual Differences, 22, 129-160.

Lovett Doust, J. W., Payne, W. D., & Podnieks, I. (1978). An ultradian rhythm of reaction time measurements in man. Neuropsychobiology, 4, 93-98.

Peterken, C., Brown, B., and Bowman, K. (1991). Predicting the future position of a moving target. Perception, 20, 5-16.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. Psychological Bulletin, 114, 510-532.

Rosenbaum, D. (1955). Perception and extrapolation of velocity and acceleration. Journal of Experimental Psychology: Human Perception and Performance, 1, 305-403.

Runeson, S. (1975). Visual prediction of collisions with natural and nonnatural motion functions. Perception and Psychophysics, 18, 261-266.

Thurstone, L. L. (1938). Primary mental abilities. Chicago: University of Chicago Press.

THE ONSET OF IRREVERSIBLE CIRCULATORY SHOCK INDUCED BY MILLIMETER WAVE EXPOSURE IN RATS

Maria Tehrany
Research Associate
Department of Biology

Trinity University
715 Stadium Dr.
San Antonio, TX 78212-7200

Final Report for :
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by :
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1995

THE ONSET OF CIRCULATORY SHOCK INDUCED BY MILLIMETER WAVE EXPOSURE IN RATS

Maria Tehrany
Research Associate
Department of Biology
Trinity University

ABSTRACT

Exposure to 35-GHZ microwave heating leads to circulatory shock and subsequent death. The purpose of this study was to determine at what mean arterial pressure (MAP) circulatory shock is irreversible. Four groups of rats (n =6 groups) were anesthetized with ketamine, surgically implanted with a catheter in the left carotid to measure MAP and exposed to 35 GHZ radiofrequency radiation until MAP was equal or less than 70 mmHg, 80 mmHg, 90 mmHg or control (MAP matched pre-exposed MAP). During exposure, EKG, temperatures at five sites and MAP was measured. Rats in 90 mmHg and control groups, for the most part, survived after MMW radiation, while most rats in 70 mmHg and 80 mmHg did not. Concluding, there is irreversible circulatory failure in rats, in most instances, after MMW exposure until MAP$\leq$ 80 mmHg.

THE ONSET OF IRREVERSIBLE CIRCULATORY SHOCK INDUCED BY MILLIMETER WAVE EXPOSURE IN RATS

Maria Tehrany

INTRODUCTION

Heat stroke may be induced by sustained exposure to environmental heating. Rats subject to moderate heating thermoregulate through sympathetic nerve responses, redistributing blood flow from visceral to cutaneous vascular beds along their tail.. Excessive heating results in hyperthermia; a rise in internal temperature is accompanied by an increase in heart rate and, initially, mean arterial pressure (MAP). If heating continues, heat stroke ensues and is characterized by a sudden loss of vasoconstriction of the mesenteric arterial bed, causing MAP to fall (Kregel et al, 1988). A hypotensive state develops, leading to circulatory shock and, if left untreated, death. The exact physiological mechanisms producing this heat induced loss of compensatory vasoconstriction is unknown. It has been shown that it is not due to decreases in either sympathetic nerve responses or catecholamines (Kregel et al. 1990).

Previously it has been shown that sustained exposure to radiofrequency radiation of millimeter wavelength (MMW) also produces hyperthermia, hypotension and death (Ryan et al, in press). Preliminary observations indicated that MAP will continue to decline until death MMW radiation is discontinued when MAP declines to 75 mmHg. I is then the purpose of this study to systematically determine the level of MMW exposure required to produce irreversible circulatory failure.

MATERIALS AND METHODS

Animals and Instrumentation

Twenty-four male Sprague-Dawley rats (Charles River Laboratories) weighing between 350 and 400 g (373 ± 2 g) were used in this study. Prior to experimentation, animals were housed in polycarbonate cages with free access to Purina rodent chow and water and were maintained on a 12h/12h light/dark cycle (lights on at 0600) in a climatically controlled room (ambient temperature = 24.0 ± 0.5°C).

Ketamine-HCl (Vetalar; 150 mg/kg, i.m.) was initially administered with supplemental doses given

during experimentation. Ketamine administration at this dose level has been shown to produce prolonged surgical anesthesia in Sprague-Dawley rats (Jauchem and Frei, 1994; Frei and Jauchem, 1989). A Teflon catheter was placed in the carotid artery for measurement of arterial blood pressure. This catheter was attached to a pre-calibrated blood pressure transducer (Century, Model CP-01) that was connected to a pressure processor (Gould,Model 13-4615-52). A lead II ECG was obtained by use of nylon-covered fluorocarbon leads attached to a Gould ECG/Biotach amplifier (Gould, Model 20-4615-65). All measured variables were recorded continuously throughout experimentation on a Gould TA2000 recorder.

The animals were also instrumented to monitor temperature at five sites: (1) left subcutaneous (lateral, mid-thoracic, side facing the MMW antenna; T_{sl}); (2) right tympanic (T_{tyr}); (3) left tympanic (T_{tyl}); (4) colonic (5-6 cm post-anus; T_c); and, (5) tail (subcutaneous, dorsal, 1 cm from base; T_{ta}). All temperature measurements were obtained via thermistor probes (BSD Medical Corporation, Model BSD-200). These temperature probes have a conductivity similar to biological tissue, are biologically inert, and produce no alterations in either the radiofrequency radiation field or in tissue energy absorption rate (Bowman, 1976). All temperature and cardiovascular data were A/D (analog-to-digital) converted by an IBM-compatible custom-designed Physiological Monitoring System (Berger *et al.*, 1991) with graphics display and data analysis capabilities.

MMW Equipment

Continuous-wave 35-GHz fields were generated by a Millimeter Wave Exposure System (Applied Electromagnetics, Inc.)and were transmitted through a Model 3-28-725 standard horn antenna (MaCom). Irradiation was conducted under far-fields conditions with the animal centered along the boresight, 110 cm from the antenna. The incident power density of the field was determined at the exposure site with an electromagnetic radiation monitor (Model 8616, Narda Microwave Corporation, employing a Model 32029 probe). The generator power output was monitored throughout exposure with a Model 4-32-B power meter (Hewlett-Packard). Irradiation was conducted in an Eccosorb RF-shielded anechoic chamber (Applied Electromagnetics, Inc.) at the Radiofrequency Radiation Facility at the USAF Armstrong Laboratory, Brooks Air Force Base, TX. Chamber temperature and relative humidity were maintained at 27.0 ± 0.5°C and 20 ± 5%, respectively, during experimentation.

Experimental Procedures

In preliminary experiments it was observed that discontinuing irradiation at a mean arterial blood pressure (MAP) of 75 mmHg caused a continuous decline in MAP to death. This protocol was designed to systematically test these observations of circulatory failure in rats to determine at what point circulatory failure is irreversible. Four groups of rats (n=6 group) were continually irradiated until MAP reached 70 mmHg, 80 mmHg, 90 mmHg or the pre-MMW baseline level (MAP before radiofrequency radiation).

Rats were anesthetized with ketamine and surgically prepared. After surgery, the rat was placed on a holder consisting of seven 0.5-cm (O.D.) Plexiglas rods mounted in a semicircle pattern between two Plexiglas plates (0.5 cm thick).After a 5 minute control period, the animal was exposed in E orientation (left lateral exposure, long axis of body parallel to electric field) to 35-GHz radio frequency radiation at an incident power density of 75mW/cm^2. This power density resulted in a whole-body specific absorption rate (SAR) of 13.0 W/kg. SARs were determined calorimetrically on six rat carcasses according to the method of Padilla and Bixby (1986).

After MMW exposure the rat was continually monitored, with supplemental doses of ketamine every half hour (if necessary), until death or for a maximum time of two hours, at which point it was euthanized. Cardiovascular and temperature parameters were recorded throughout MMW exposure and after cessation of MMW exposure.

Data Analysis

A one way Analysis of Variance (ANOVA) was use to determine differences in temperatures, heart rate and MAP before and at cessation of irradiation. If significance was obtained, ($p<0.05$) the Student-Newman-Keuls (SNK) multiple comparison test followed to determine differences between the means. Means ± SE (standard error) were calculated for heart rate and MAP at intervals of five minutes, to compare irradiation and survival periods (the period following RFR) for 70 mmHg, 80 mmHg, 90 mmHg and control exposed rats. The 5 minute intervals extended to the mean irradiation and survival time.

RESULTS

Among the experimental groups, there were no significant differences between control temperatures

(p>0.05). All temperatures significantly increased from time of control to the end of MMW exposure, noting that left skin (T_{sl}) had the greatest absolute increase (table 1). At the cessation of irradiation the colonic temperature (T_c) for 80 mmHg was shown to be significantly different from 70 mmHg, 90 mmHg and control groups, and the T_{sl} showed significance between all groups.

Table 1. The mean temperature (n=6) ± standard error, at the five sites, for each group right before RF exposure (control) and at the end of RF exposure.

	Control Temperature (°C)				
Experimental Group	**colonic (T_c)**	**right tympanic (T_{tyr})**	**left tympanic (T_{tyl})**	**tail (T_{ta})**	**left skin (T_{sl})**
70 mmHg	36.8 ± 0.1	39.9 ± 0.2	36.7 ± 0.1	29.5 ± 0.4	34.6 ± 0.6
80 mmHg	37.1 ± 0.1	36.8 ± 0.2	36.6 ± 0.2	30.2 ± 0.5	34.9 ± 0.5
90 mmHg	36.8 ± 0.2	36.5 ± 0.1	36.3 ± 0.1	29.5 ± 0.5	34.5 ± 0.5
control	36.9 ± 0.2	36.8 ± 0.2	36.7 ± 0.2	29.8 ± 0.5	34.9 ± 0.4
	RF Cessation Temperature (°C)				
Experimental Group	**colonic (T_c)**	**right tympanic (T_{tyr})**	**left tympanic (T_{tyl})**	**tail (T_{ta})**	**left skin (T_{sl})**
70 mmHg	39.7 ± 0.4	40.4 ± 0.5	40.3 ± 0.4	32.7 ± 0.7	46.9 ± 0.5
80 mmHg	40.8 ± 0.2	40.0 ± 0.2	40.0 ± 0.1	33.9 ± 0.6	46.0 ± 1.0
90 mmHg	39.7 ± 0.3	39.5 ± 0.4	39.4 ± 0.3	31.6 ± 0.9	45.3 ± 0.9
control	39.2 ± 0.3	39.3 ± 0.2	39.6 ± 0.2	33.6 ± 0.4	45.8 ± 0.2

Ten minutes after the onset of MMW exposure blood pressure peaked for all groups then continued on in a gradual decline (figure 1). MMW exposure also produced a gradual and steady increase in heart rate (HR) in all groups (figure 2)

After the cessation of irradiation blood pressure dropped more drastically (figure3). The control group decreased the slightest, leveling out quickly; all of these rats survived for the 120 minutes observation period.

Comparatively 90 mmHg and 80 mmHg groups had the largest decrease in blood pressure. The curve for 90 mmHg shows an increase that levels out to almost the same point as control but discontinues at a mean time of 85 minutes where 1 rat survived 22 minutes and 5 rats survived the full 120 minutes after MMW exposure. Where as, the curves for 80 mmHg and 70 mmHg groups cease at survival times of 50 minutes (1 rat survived 120 min after RFR) and 15 minutes (0 rats survived 120 min after RFR) (figure 3).

The patterns of HR after the termination of irradiation diverge (figure 4).HR in rats exposed to only moderate levels of MMW (i.e. 90 mmHg and control groups) demonstrate a continuous decrease. Rats which were exposed to a more severe irradiation (to 80 mmHg and 70 mmHg) also demonstrated an overall decline in HR until death, although their values fluxuated greatly (Figure 4).

Figure 1. Blood pressure versus mean RF time for all groups during RF exposure.

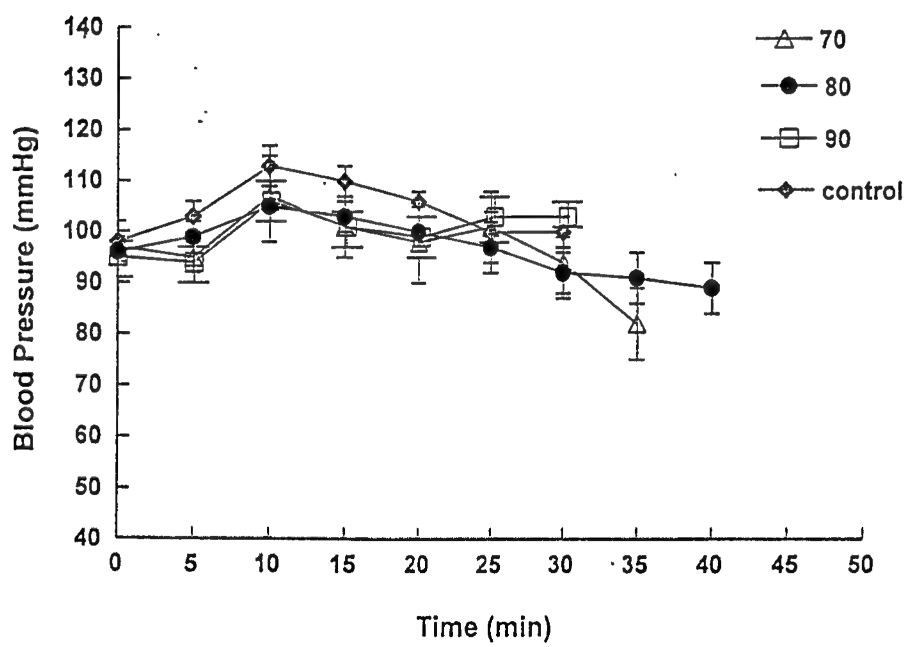


Figure 2. Heart rate versus mean RF time for all groups during RF exposure.

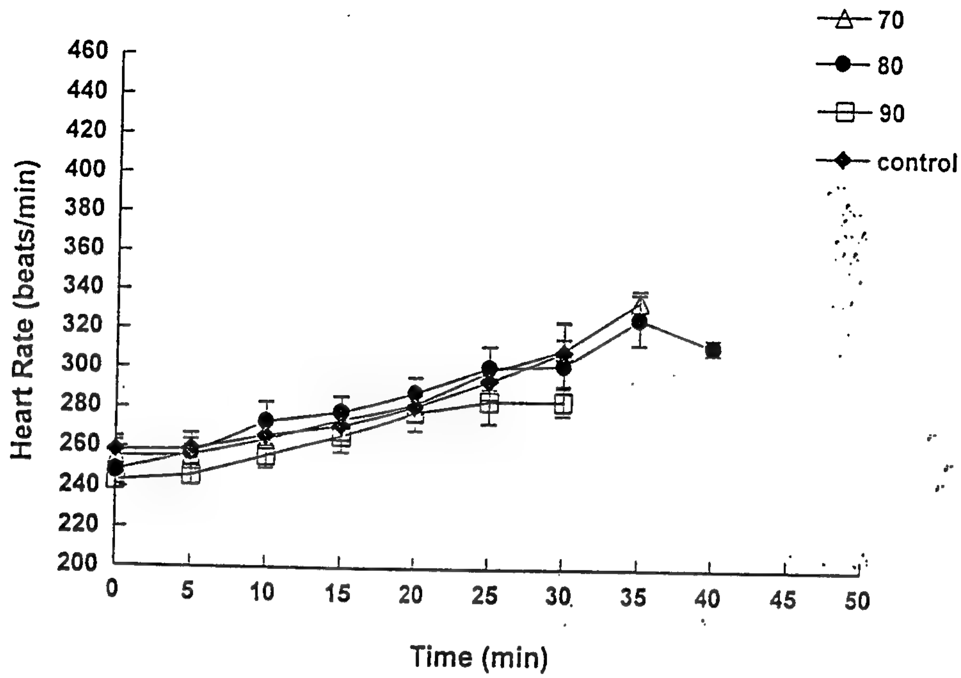


Figure 3. Blood pressure versus mean survival time for each group after cessation of RF exposure.

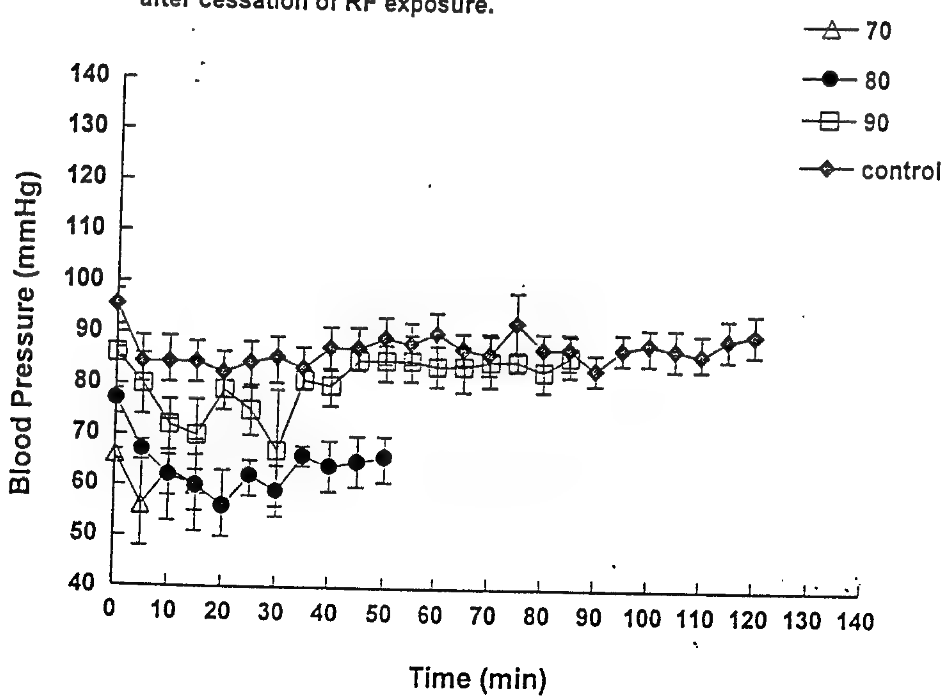
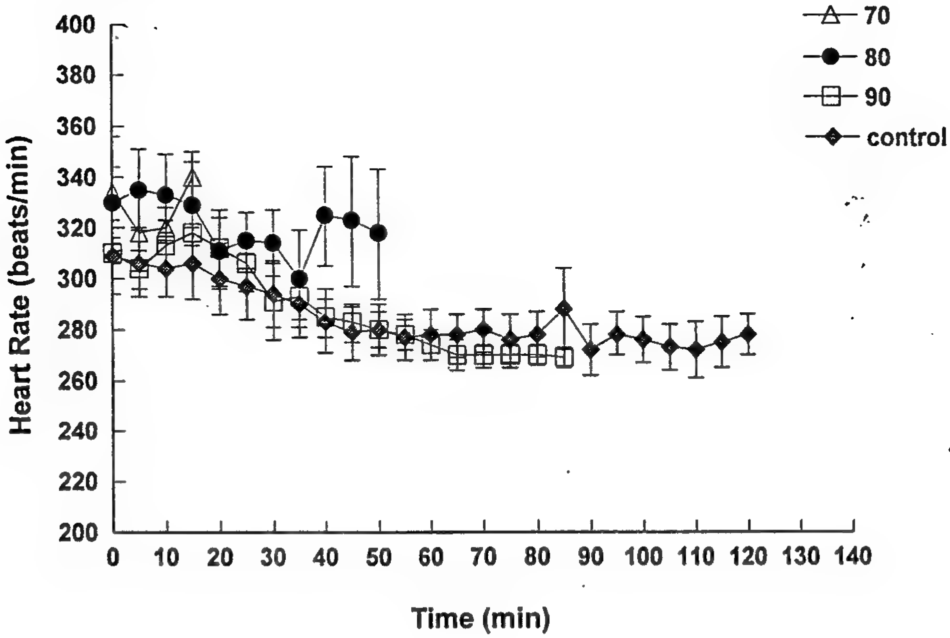


Figure 4. Heart rate versus mean survival time for each group after cessation of RF exposure.



CONCLUSION

The increase in time of exposure to MMW, between experimental groups, paralleled the increase in T_{sl}. As T_{sl} rose along with overall body temperature, thermoregulatory processes were overwhelmed until heat stroke and hypotension was apparent. It is evident from this study that the range of blood pressure for irreversible circulatory failure induced by 35 GHz MMW exposure is between 80 mmHg and 70 mmHg . When MMW exposure was discontinued at this point, hypotension could not be reversed and death occurred. Comparatively, the discontinuation of irradiation at 90 mmHg and control (at or near initial blood pressure) did not result in irreversible circulatory failure and death; MAP in both groups demonstrated a sustained recovery, although not to control levels. Additionally, the vast majority of animals which received only this mild exposure level survived for the entire 2 hour observation period . From this study, I conclude that irreversible circulatory failure occurs only when MMW exposure is continued to a point of significant hypotension (i.e., MAP equals at least 80 mmHg).

REFERENCES

Berger, R.E., Frei, M.R. and Jauchem, J.R. Computer-based monitoring of thermoregulatory responses of anesthetized rats exposed to radiofrequency radiation, *Physiologist* 34: 246, 1991.

Bowman, R.R. A probe for measuring temperature in radio-frequency heated material. *IEEE Trans. Microwave Theory Tech.* 24: 43-45, 1976.

Frei, M.R. and Jauchem, J.R. Effects of 2.8-GHz microwaves on restrained and ketamine-anesthetized rats. *Radiat. Environ. Biophys.* 28:155-164, 1989.

Gisolfi, C.V., Matthes,R.D., Kregel, K.C., and Oppliger, R. Splanchic sympathetic nerve activity and circulating catecholamines in the hyperthermic rat. *J. Appl. Physiol.* 70: 1821-1826, 1991.

Jauchem, J.R., and Frei, M.R. Cardiorespiratory changes during microwave-induced lethal heat stress and β-adrenergic blockade. *J. Appl. Physiol.* 77: 434-440, 1994.

Kielblock, A.J., Strydom, N.B., Burger, F.J., Pretorius, P.J. and Manjoo, M. Cardiovascular origins of heatstroke pathophysiology: an anesthetized rat model. *Aviat. Space Environ. Med.* 53: 171-178, 1982.

Kregal, K.C. and Gisolfi. Peripheral vascular responses to hyperthermia in the rat. *J. Appl. Physiol.* 64:2582-2588, 1988.

Kregal, K.C. and Gisolfi, C.V. Circulatory responses to vasoconstrictor agents during passive heating in the rat. *J. Appl. Physiol.* 68:1220-1227, 1990.

Padilla, J.M., and Bixby, R.R. Using Dewar-flask calorimetry and rectal temperature to determine the specific absorption rates of small rodents. Technical Report USAFSAM 86-3, US Air Force School of Aerospace Medicine, Brooks Air Force Base, TX, 1986.

Ryan, K.L., Lovelace, J.D., Frei, M.R, and Jauchem, J.R. Exogenous Nitric Oxide Does Not Affect Hypotension Induced by 35-GHz Microwave Heating. In preparation.

STATISTICS IN CREW SYSTEMS DIRECTORATE

Renardo D. Tyner
Graduate Student
Department of Discrete and
Statistical Sciences

Auburn University
Department of Discrete
and Statistical Sciences
120 Math Annex
Auburn,AL 36849-5307

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1995

STATISTICS IN CREW SYSTEMS DIRECTORATE

Renardo D. Tyner
Graduate Student
Department of Discrete and
Statistical Sciences

Abstract

The field of statistics is applied in many areas encompassing scientific research, as a result, statisticians are afforded an opportunity to be involved in the analysis of data from many different scientific experiments. To be a statistician in the Crews Systems Directorate, allows one to become a skillful practionner of the statistical tools of the trade, while performing statistical analyses within a multi-discipline structure. Since, the statisticians' job is a unique one, mere words can not express the invaluable experience I've had during the summer program. Therefore, a brief synopsis of one study will be discussed.

A retrospective study on the number of injuries caused by helmet loss after ejection from fighter aircraft was done. Researchers were interested in comparing the incidence of mishaps associated with helmet loss in the Air Force and Navy. Incidentally, the study compared data from the Air Force, Navy, Air Force and Navy (combined).

STATISTICS IN
CREW SYSTEMS DIRECTORATE

Renardo Tyner

Introduction

In recent years, there were a number of aircraft ejection seat mishaps in the Air Force and Navy. Consequently, officials wanted to know the cause of these mishaps. Therefore, a study was done to determine if there was a correlation between helmet loss and mishaps. Incidentally, there were more helmet loses associated with mishaps in the Navy than Air Force. By constructing contingency tables, we were able to determine if there was a correlation between the Air Force and Navy.

Discussion

There are three questions that were of concern to the researchers: Does helmet loss correlate to a specific airspeed? Does helmet loss correlate with head injury? Does helmet loss correlate with neck injury? Incidently, these questions were asked for the Air Force, Navy, and Air Force/Navy. For the first question, Air Force pilots typically have faster airspeeds at the time of ejection than Navy pilots. Due to this fact, it is believed that there are more Air Force pilots who eject and lose their helmets than Navy pilots. Secondly, the number of pilots who injured their head/neck resulting from helmet loss in the Air Force is likely to be more than Navy pilots. During faster airspeeds, which is typical in the

Air Force, the wind gets under the helmets and causes it be lifted off the head resulting in head/neck injury.

Results

The Chi-square test was performed to determine the incidence of injuries associated with helmet loss. A Chi-square statistic tests if the differences in helmet loss associated with mishaps are due to chance, indicated by (prob). The results indicate that there is a correlation between neck injury and helmet loss in the Air Force(Prob=.003). Among the number of neck injuries in the Navy, statistical test indicate that there is no correlation between neck injury and helmet loss(Prob=.5111). At the same time, there is a correlation between head injury and helmet loss for both the Air Force and the Navy(Prob=.001). By comparing the data from the Air Force and Navy, there appears not to be correlation between the services and helmet loss.

Conclusion

The proportion of helmet loss is nearly the same in the Air Force and Navy. However, helmet loss caused more neck injury in the Air Force than in the Navy. Whereas, helmet loss contributed to head injury in both services.